SECTION 7

DATA EVALUATION

7.1 Introduction

7.1.1 One of the major concerns of USEPA, other federal, state and private agencies is to describe water quality and habitat quality in terms which are easily understood by the non-biologist. The purpose of this section is not to recommend one particular data evaluation method, but to point out a number of more common methods. Some of these methods may not be applicable to every stream or water body in the United States.

7.1.2 Water quality and habitat quality are reflected in the species composition and diversity, population density and biomass, and physiological condition of indigenous communities of aquatic organisms. A number of data interpretation methods have been developed based on these community characteristics to indicate the water quality and the degree of habitat degradation, and also to simplify communication problems regarding management decisions.

7.2 Analyses of Qualitative Data

7.2.1 As previously defined, qualitative data result from samples collected in such a manner that no estimates of numerical abundance or biomass can be calculated. The principle output is a list of taxa collected in the various habitats of the environment studied. The numerous schemes advanced for the analysis of qualitative data may be grouped under two categories; the indicator organism scheme and reference station methods.

7.2.2 Indicator Organism Scheme

7.2.2.1 For this technique, individual taxa are classified on the basis of their tolerance or intolerance to various levels of domestic wastes (Beck, 1954; Lewis, 1974; Chutter, 1972; Hilsenhoff, 1977; Howmiller and Scott, 1977; Milbrink, 1983; Reynoldson <u>et</u> <u>al</u>. 1989). Taxa are classified as tolerant or intolerant according to their presence or absence in different environments as determined by field studies. Beck (1955), reduced data, based on the presence or absence of indicator organisms, to a simple numerical form for ease in presentation. Clean water taxa are given twice the weight as tolerant organisms in the formula:

$$2\ (n\ \text{Class I}) + (n\ \text{Class II}) = \text{Biotic Index}$$

where "n" is the number of taxa in that class. Values less than 10 are considered to indicate a polluted stream.

7.2.3 Reference Station Methods

7.2.3.1 Reference station methods (Ohio EPA, 1989) compare the characteristics of the fauna in clean water habitats with those of fauna in habitats subject to stress. Patrick (1950) compared stations on the basis of richness of species, and Wurtz (1955) used indicator organisms in comparing stations.

7.2.4 If adequate background data are available to an experienced investigator, both of these techniques can prove quite useful; particularly for demonstrating the effects of gross to moderate organic contamination on the macroinvertebrate community. To detect more subtle changes in the macroinvertebrate community, quantitative data on numbers or biomass of organisms are needed. Data on the presence of tolerant and intolerant taxa and richness of species may be effectively summarized for evaluation and presentation by means of line graphs, bar graphs, pie diagrams, histograms, or pictorial diagrams (Ingram and Bartsch, 1960).

7.2.5 Classification of representative macroinvertebrates according to their tolerance of organic wastes is presented in Appendix A. Hilsenhoff's (1977) original tolerance classification with a numerical range of 0 to 5 is followed in Appendix A. Later, Hilsenhoff (1987) modified his biotic index for Wisconsin taxa to include more intermediate values with a numerical ranged of 0-10. However, similar results can be obtained using index values of either 0-5 or 0-10, and adequate information is not available for many species that would allow use of the more definitive 0-10 tolerance range (Hilsenhoff, 1990, personal communication). In most cases, the taxonomic nomenclature used is that of the original authors listed at the end of Appendix A. The pollutional classifications were arbitrarily placed in three categories--tolerant, facultative, and intolerant--defined as follows:

- Tolerant: Organisms frequently associated with gross organic contamination, that are generally capable of thriving under anaerobic conditions. Tolerance values 4 and 5.

- Facultative: Organisms having a wide range of tolerance that frequently are associated with moderate levels of organic contamination. Tolerance values 2 and 3.

- Intolerant: Organisms that are usually not found associated with organic contaminants and are generally intolerant of even moderate reductions in dissolved oxygen. Tolerance values 0 and 1.

When evaluating qualitative data in terms of material such as that contained in Appendix A, the investigator should keep in mind the pitfalls mentioned earlier, as well as the following:

7.2.5.1 Since tolerant species may be found in both clean and degraded habitats, a simple record of their presence or absence is not of significance. However, the presence of intolerant organisms provides evidence of only one condition--clean water. But the fact that sensitive (intolerant) species may be totally absent, because of the discharge of toxic substances or thermal pollution, would indicate that absence of intolerant species may not be a reflection of the presence of organic wastes. The presence of tolerant organisms is a significant indicator of organic pollution only when they are dominant in the sample.

7.2.5.2 The presence or absence of particular taxa may depend more on characteristics of the environment, such as velocity and substrate, than on the level of degradation by organic wastes. This affects both the original placement of the taxa in the classificatory scheme and its presence in study samples.

7.2.5.3 Because indicator species evaluations are based on the presence or absence of organisms, a single specimen has as much weight as a large population. Therefore, studies may be biased by the drift of organisms into the study area. The technique is totally subjective and dependent upon the skill and experience of the individual who makes the field collections. Therefore, results of one investigator are difficult to compare with those of another, particularly where data are summarized in an index such as that proposed by Beck (1955).

7.2.6 Biotic Index

7.2.6.1 Many of the problems discussed above can be overcome by use of the biotic index proposed by Chutter (1972) and modified by Hilsenhoff (1977) for use with the index values given in Appendix A. Any organisms not listed in Appendix A should be given an index of three (3) unless available information would suggest a different value. This same formula is used with the family level biotic index of Hilsenhoff (1988a) and the Rapid Bioassessment metric 2 of Protocol III (Plafkin et al., 1989) where pollution tolerance values of 0-10 are used. Appendix B gives the family level index values (Hilsenhoff, 1988a) for use with the family level biotic index. Results are comparable between stations in the same and nearby streams if similar habitats were sampled using similar methods and sampling effort (Hilsenhoff, 1988a,b). The formula to use is:

$$HBI = \Sigma \frac{n_i \; a_i}{N}$$

Where "n_i" is the number of individuals in the "i^{th}" taxa, "a_i" is the index value of that taxa, and "N" is the total number of individuals in the sample. Biotic index values below 1.75 indicate excellent water quality, 1.76-2.50 indicate good water quality, 2.51-3.75 indicate fair water quality, 3.76-4.00 indicate poor water quality, and over 4.00 would indicate serious water quality problems.

7.2.6.2 The following are water quality values for Hilsenhoff's (1988a) family level biotic index: 0.00-3.75 (excellent), 3.76-4.25 (very good), 4.25-5.00 (good), 5.01-5.75 (fair), 5.76-6.50 (fairly poor), 6.51-7.25 (poor), and 7.26-10.00 (very poor).

7.3 Analyses of Semi-quantitative and Quantitative Data

7.3.1 The high variability usually associated with benthic macroinvertebate populations makes them difficult to study quantitatively because of the large number of samples needed to obtain normal levels of precision. For most benthic studies, it is generally impractical, due to large number of samples needed, to detect population changes of less than 100% of the mean. Many benthic populations exhibit such high variability (see Section 4.5.) that any reasonable number of replicate samples would be too small to detect a population density difference of more than 200% of the mean between two sites (Schwenneker and Hellenthal, 1984). It is important to keep this limitation in mind as one considers the methods to use in evaluating the data.

7.3.2 Data from quantitative samples may be used to obtain total standing crop of individuals, or biomass, or both and numbers or biomass, or both, of individual taxa per unit area or unit volume or sample unit. Data from quantitative samples may also be evaluated in the same manner as discussed for qualitative samples but results will be qualitative. In order to reduce the amount of time spent in field sampling, there has been a recent trend to collect data based on level of effort or other not strictly quantitative methods and treat the data as semi-quantitative. These data are then analyzed using the quantitative methods described in this section.

7.3.3 For purposes of comparison and to provide data useful for determining production, a uniform convention must be established for the units of data reported. For this purpose, USEPA biologists should adhere to the following units:

- Data from devices sampling a unit area of bottom are reported in grams dry weight or ash-free dry weight per square meter (gm/m^2), or numbers of individuals per square meter, or both.
- Data from multiplate samplers are reported in terms of the total surface area of the plates, as grams dry weight or ash-free dry weight or numbers of individuals per square meter, or both.
- Data from rock-filled basket samplers are reported as grams dry weight, ash-free dry weight, or numbers of individuals per sampler, or both.

7.3.4 Three informative parameters of benthic community structure which may be obtained from quantitative grab or artificial substrate sample data are standing crop (biomass or numbers), species richness, and species composition. Standing crop and species richness in a community are highly sensitive to natural environmental conditions and to anthropogenic perturbations resulting from the introduction of contaminants. These parameters, particularly standing crop, may vary considerably in unpolluted habitats, where they may range from the typically high standing crop of littoral zones of glacial lakes to the sparse fauna of torrential soft-water streams. Thus, it is important that comparisons be made only between truly comparable habitats. Typical responses of standing crop or species richness to various types of stress are shown in Table 7 below:

7.3.5 Organic enrichment and sludge deposits are frequently associated. The responses shown are by no means simple or fixed and may vary depending on a number of factors including a combination of stresses acting together or in opposition, indirect effects (such as the destruction of highly productive vegetative substrate by temperature alterations, sludge deposits, turbidity, or chemical weed control) and the physical characteristics of the stressed environment; particularly in relation to substrate and current velocity.

Table 7. TYPICAL RESPONSES TO VARIOUS TYPES OF STRESS BY PARAMETERS OF BENTHIC COMMUNITY STRUCTURE

Stress	Standing crop (Numbers or Biomass)	Number of Taxa
Toxic substance	Reduces	Reduces
Severe temperature changes	Variable	Reduces
Silt	Reduces	Reduces
Low pH	Reduces	Reduces
Inorganic nutrients	Increases	Variable
Organic enrichment (Low DO)	Increases	Reduces
Sludge deposits (Non toxic)	Increases	Reduces

7.3.6 Data on standing crop and species richness may be presented in simple tabular form or pictorially with bar and line graphs, pie diagrams, and histograms. Whatever the method of presentation, the number of replicates and the sampling variability must be shown in the tables or graphs. Sampling variability may be shown as a range of values or as a calculated standard deviation, as discussed in Section 7.6.

7.3.7 Data on standing crop and species richness are amenable to simple but powerful statistical techniques of evaluation. Under grossly stressed situations, such analyses may be unnecessary; however, in some cases, the effects of environmental perturbations may be so subtle in comparison with sampling variation that statistical comparisons are a helpful and necessary tool for the evaluation process. For this purpose, biologists engaged in studies of macroinvertebrates should familiarize themselves with the simple statistical tools discussed in Section 7.6.

7.3.8 The usefulness of species composition as a parameter of environmental quality is based on the generally observed phenomenon that relatively undisturbed environments support communities having large numbers of species with no individual species present in overwhelming abundance. If the species found in a random sample from such a community are ranked on the basis of their numerical abundance, there will be relatively few species with large numbers of individuals and large numbers of species represented by only a few individuals. Many forms of stress alter species composition by making the environment unsuitable for some species or by giving other species a competitive advantage.

7.3.9 It is important for the investigator to keep in mind that there are naturally occurring severely stressed environments supporting communities dominated by one or more species adapted to rigorous conditions. Examples include the profundal fauna of deep lakes and the black fly dominated communities of the high gradient, bedrock section of a torrential stream. Furthermore, because colonization is by chance, both species richness and species composition may be highly variable in a successional community; for this reason, data summarized from artificial substrate samples must be evaluated with caution. These confounding factors can be reduced by comparing data from similar environments and by exposing artificial substrate samplers long enough for a relatively stable community to develop.

7.3.10 Data on species composition may be summarized and evaluated using percent species composition tables, frequency distribution tables and/or graphs; however, for any appreciable number of samples, such methods of presentation are so voluminous that they are virtually impossible to compare and interpret. Fortunately, single numerical values which provide a measure of species composition can be extracted from indices of diversity as proposed by Margalef (1957) and subsequently utilized by numerous workers (McIntosh, 1967; Cairns and Dickson, 1971; Wilhm and Dorris, 1968). Mean diversity ($\bar{d}$) may be calculated using the machine formula presented by Lloyd, Zar, and Karr (1968) and better known as the Shannon-Weaver mean diversity (Shannon and Weaver, 1963).

$$\bar{d}=\frac{C}{N}\left(N \log_{10} N - \Sigma n_i \log_{10} n_i\right)$$

where C=3.321928 (converts base 10 log to base 2); N = total number of individuals; and n_i = total number of individuals in the i^{th} species. When their table (see Table 23) is used, the calculations are simple and straightforward, as shown in Table 8.

Table 8 EXAMPLE OF CALCULATION OF MEAN DIVERSITY

	Taxa Number	Number of Individuals in each Taxon (n_i)	$n_i \log n_i$ (From Table 23)
	1	41	66.1241
	2	5	3.4949
	3	18	22.5949
	4	3	.4314
	5	1	.000
	6	22	29.5333
	7	1	.0000
	8	2	.6021
	9	12	12.9502
	10	4	2.4082
Totals	10	109	139.1391

N log N (109) = 222.0795 (From Table 23)

$\Sigma n_i \log_{10} n_i$ = 139.1391 (From Column 3 above)

$$\bar{d} = \frac{3.321928}{109}\ (222.0795-139.1391)$$

$$\bar{d} = 0.030476 \times 82.9404$$

$$\bar{d} = 2.5$$

7.3.10.1 Mean diversity as calculated above is affected both by richness of species and by the distribution of individuals among the species (species composition) and may range from zero to 3.321928 log N. Since the calculated value of mean diversity is a result of the interaction of two parameters which may vary independently, it is often insensitive to subtle changes in community structure. Therefore, unless the environment has been grossly modified, mean diversity ($\bar{d}$) often has limited value in detecting alterations in community structure and serves mainly as an intermediate step in the calculation of a single numerical value for species composition.

7.3.11 To evaluate the component of diversity due to the distribution of individuals among the species (species composition), the calculated $\bar{d}$ must be compared with a hypothetical maximum $\bar{d}$ based on an arbitrarily selected distribution. The measure of redundancy proposed by Margalef (1957) is based on the ratio between $\bar{d}$ and a hypothetical maximum computed as though all species were equally abundant. In nature, equality of species is quite unlikely, so Lloyd and Ghelardi (1964), proposed the term "equitability" and compared $\bar{d}$ with a maximum based on the distribution obtained from MacArthur's (1957) broken stick model. The MacArthur model results in a distribution quite frequently observed in nature; one with a few relatively abundant species and increasing numbers of species represented by only a few individuals. Sample data are not expected to conform to the MacArthur model, since it is only being used as a yardstick against which the distribution of abundances is being compared. Lloyd and Ghelardi (1964) devised a table for determining equitability by comparing the number of species (s) in the sample with the number of species expected (s′) from a community that conforms to the MacArthur model. In the table (reproduced as Table 24) the proposed measure of equitability is:

$$e = \frac{s'}{s}$$

where s = the number of taxa in the sample and s′ = the tabulated value.

7.3.11.1 For the example given above:

$$e = \frac{s'}{s} = \frac{8}{10} = 0.8$$

where "s′" is found from Table 24 using $\bar{d}$ of 2.5. Equitability "e", as calculated, may range from 0 to 1 except in the unusual situation where the distribution in the sample is more equitable than the distribution resulting from the MacArthur model. Such an eventuality will result in values of "e" greater than 1, and this occasionally occurs in samples containing only a few specimens with several taxa represented. The value of "e" is not entirely sample size independent and should not be used for samples containing fewer than five taxa.

7.3.11.2 Equitability ("e") is very sensitive to slight changes in community structure. Since the sample is a representation of the community sampled, a usable index must be sensitive to sample differences and within station variability must be handled by proper study design and adequate replication. Equitability above 0.5 is indicative of waters not affected by oxygen demand wastes. Even slight levels of degradation have been found to reduce equitability below 0.5, generally below 0.3.

7.3.12 Quantitative data can also be produced using the biotic index described in 7.2.6 as long as quantitative methods were used in sample collection and analysis, and proper assumptions are made concerning the subjective nature of the pollution tolerance values.

7.3.13 A rather simple technique for evaluating quantitative data is the sequential comparison index (SCI) which estimates relative differences in biological diversity (Cairns and Dickson, 1971). The method requires no taxonomic expertise on the part of the investigator and is based on differences in the shape, color, and size of the organisms. It should be stressed that the method is useful only as a technique to evaluate the diversity of the bottom community rapidly producing numerical data which can be interpreted statistically. However, it should not be used to replace other more exact techniques providing information on the identity and pollution tolerance of the organisms and requiring persons trained in aquatic ecology.

7.3.14 Wilhm's Species Diversity Index (Wilhm and Dorris, 1968) is based upon information theory and is an attempt to give a numerical value to the environmental changes caused by waste dischargers. This index takes into account not only the number of species encountered, but also the relative abundances of the different species and is very similar to that described in section 7.3.10. Results from this system indicate that values of "d" less than one are indicative of heavy pollution, values from one to three indicate moderate pollution and values above three are found in clean water areas.

7.3.15 Harkins and Austin (1973) have also developed a method that appears to be universal in scope and has worked well in diverse situations. This method is based on average diversity per individual and redundancy which are reduced to a single index value per sample utilizing a nonparametric discrimination technique which then gives a unique distance value from a predefined "biological desert" condition (control values). This condition exists as the case of no organisms present or only one species containing "n" organisms.

7.3.15.1 Computer programs have been written to perform the needed calculations as well as the analysis of variance which can be used with this method. Harkin and Austin's method then is essentially an objective method for reducing several biological indexes to a single meaningful value that will reflect subtle changes in the structure of aquatic communities. The resulting sets of standardized distance values can be compared subjectively or can be subjected to statistical evaluation and probability level of differences assessed. With this method any changes of quality will be detected and can be plotted for long-term trend analysis.

7.4 Rapid Bioassessment Techniques

7.4.1 Rapid Bioassessment Techniques (Plafkin et al., 1989) are generally considered both qualitative and semi-quantitative. The protocols were established as a rapid means of detecting aquatic life impairments and assessing their relative severity and are not intended to replace traditional biomonitoring methods. The three protocols each consist of three basic components: water quality/physical characteristics, habitat assessment, and a biosurvey. The biological assessment in each protocol involves an integrated analysis of both functional and structural components of the aquatic communities through use of metrics for benthic macroinvertebrates and fish.

7.4.1.1 Rapid Bioassessment Protocol I consists of an estimation of the level of diversity of the aquatic biota; an estimation of the relative abundance of major macrobenthic taxa, using a qualitative sampling process to include as many habitats as possible; observations of the presence of fish, plants and physical structures; observations on habitat alterations; and observation on possible sources of impact.

7.4.1.2 Rapid Bioassessment Protocol II consists of an in the field estimation of the abundance level of the major aquatic biota, a list of families found in a 100-organisms subsample based on field identification, the number of individuals in each family, and separation of these into scraper and filtering collector functional feeding groups, collection of a course particulate organic material (CPOM) sample, and observations as in Protocol I.

7.4.1.3 Rapid bioassessment Protocol III is similar to Protocol II except that the subsampling and identifications are done in the laboratory and the organisms are identified to genus or species.

7.4.1.4 Rapid Bioassessment Protocols IV and V are based on fish surveys conducted by fishery personnel usually with assistance from the aquatic biologist involved with Protocols I to III.

7.5 Community Metrics and Pollution Indicators

7.5.1 Biological impairment of the benthic community may be assessed by use of metrics including community, population and functional parameters. Metrics measure different components of the community structure and have different ranges of sensitivity to stress. It is advisable, therefore, to use several metrics because an integrated approach provides more assurance of a valid assessment. A few of the more useful metrics are briefly described.

7.5.2 Species (or Taxa) Richness reflects the health of the community through a measurement of the variety of taxa (total number of families and/or genera and/or species) present. Richness generally increases with increasing water quality, habitat diversity, and/or habitat suitability. Sampling of highly similar habitats will reduce the variability in this metric attributable to factors such as current speed and substrate type. Some pristine headwater streams may be naturally unproductive, supporting only a very limited number of taxa. In these situations, organic enrichment may result in an increase in number of taxa.

7.5.3 The modified Hilsenhoff Biotic Index (HBI) (Plafkin et al. 1989) was developed to summarize overall pollution tolerance of the benthic arthropod community with a single value. This index was developed as a means of detecting organic pollution in communities inhabiting rock or gravel riffles/runs. Although Hilsenhoff's (1977) biotic index using tolerance values of 0-5 was originally developed for use in Wisconsin, it is successfully used by several states and should prove reliable for extensive use, perhaps requiring regional modification in some instances. Based on an in depth study of 53 Wisconsin streams Hilsenhoff (1988a) expanded the scale for tolerance values to 0-10. The 0-10 scale was adopted for use with the Rapid Bioassessment Protocol III and was modified to include non-arthropod species.

7.5.3.1 Although it may be applicable for other types of pollutants, use of the HBI in detecting non-organic pollution effects has not been thoroughly evaluated. The state of Wisconsin is conducting a study to evaluate the ability of Hilsenhoff's index to detect non-organic effects. Winget and Mangum (1979) have developed a tolerance classification system applicable to the assessment of nonpoint source impact.

7.5.3.2 Invertebrate Community Index (ICI)--Ohio EPA (1989) measures the condition of the macroinvertebrate community by use of the Invertebrate Community Index (ICI). This index is a modification of the Index of Biotic Integrity (IBI) used for fish (Karr, 1981) consisting of ten community metrics. Scoring of each metric varies with drainage area and ecoregion (Ohio EPA, 1987), and all but one metric is generated from artificial substrate (multiplate) samplers. Metric 10 is based solely on qualitative sample data.

7.5.4 Ratio of Scraper and Filtering Collector Functional Feeding Groups reflect the riffle/run community food base and provides insight into the nature of potential disturbance factors. The proportion of the two feeding groups is important because predominance of a particular feeding type may indicate an unbalanced community responding to an overabundance of a particular food source. The predominant feeding strategy reflects the type of impact detected.

7.5.4.1 A description of the functional feeding group concept can be found in Cummins (1973). Genus-level functional feeding group designations for most aquatic insects can be found in Merritt and Cummins (1984). Within a functional feeding group individual taxa may be either specialists which are restricted to the utilization of a specific food resource or be facultative and thus be able to exploit a broader range of food resources. The trophic generalists (see Merritt and Cummins, 1984) are expected to be better able to tolerate disturbance to aquatic habitats and thus become numerically dominant because of their more flexible ability to utilize available resources.

7.5.4.2 The relative abundance of scrapers and filtering collectors in the riffle/run habitat provides an indication of the periphyton community composition and availability of suspended fine particulate organic material (FPOM) associated with organic enrichment. Scrapers increase with increased abundance of diatoms and decrease as filamentous algae and aquatic mosses (which cannot be efficiently harvested by scrapers) increase. However, filamentous algae and aquatic mosses provide good attachment sites for filtering collectors, and the organic enrichment often responsible for over abundance of filamentous algae provide FPOM utilized by the filterers.

7.5.4.3 Filtering collectors are also sensitive to toxicants bound to fine particles and may decrease in abundance when exposed to sources of such bound toxicants. The scraper-to-filtering-collector ratio may not be a good indication of organic enrichment if adsorbing toxicants are present. This situation is often associated with point source discharges where certain toxicants adsorb readily to dissolved organic matter forming FPOM during flocculation. Toxicants thus become available to filterers via FPOM.

7.5.5 <u>Ratio of Shredder Functional Feeding Group and Total Number of Individuals</u> collected in a coarse particulate organic material (CPOM) sample is also based on the functional feeding group concept. The abundance of the shredder functional group relative to the abundance of all other functional groups allows evaluation of potential impairment as indicated by the CPOM-based shredder community. Shredders are sensitive to riparian zone impacts and are particularly good indicators of toxic effects when the toxicants involved are readily adsorbed to the CPOM and either affect the microbial communities colonizing the CPOM or the shredders directly (Plafkin <u>et al</u>. 1989).

7.5.5.1 The degree a toxicant effects shredders versus filterers depends on the nature of the toxicant and the organic particle adsorption efficiency. Generally, as the size of the particle decreases, the adsorption efficiency increases as a function of the increased surface to volume ratio (Hargrove 1972). Toxicants of a terrestrial source (pesticides and herbicides) accumulate on CPOM prior to leaf fall thus having a substantial effect on shredders. The focus of this approach is on a comparison to the reference community, which should have an abundance and diversity of shredders representative of the particular area under study. This allows for an examination of shredder or collector "relative" abundance as indicators of toxicity.

7.5.6 <u>Ratio of Ephemeroptera-Plecoptera-Trichoptera (EPT) and Chironomidae</u> abundance uses relative abundance of these indicator groups as a measure of community balance. Good biotic condition is reflected in communities having a fairly even distribution among all four major groups and with substantial representation in the sensitive groups Ephemeroptera, Plecoptera, and Trichoptera. Skewed populations having a disproportionate number of the generally tolerant Chironomidae relative to the more sensitive insect groups may indicate environmental stress (Ferrington 1987). Certain species of some genera such as Cricotopus are highly tolerant (Lenat, 1983; Mount et al., 1984), opportunistic,and may become numerically dominant in habitats exposed to metal discharges where EPT taxa are not abundant, thereby providing a good indicator of toxicant stress (Winner et al., 1980; Clements et al., 1988).

7.5.6.1 Chironomids tend to become increasingly dominant in terms of percent taxonomic composition and relative abundance along a gradient of increasing enrichment or heavy metals concentration (Ferrington 1987).

7.5.7 The EPT Index (the total number of distinct taxa within the orders Ephemeroptera, Plecoptera, and Trichoptera) compared to total taxa present generally increases with increasing water quality. This value summarizes taxa richness within the insect orders that are generally considered to be pollution sensitive. Headwater streams which are naturally unproductive may experience an increase in taxa (including EPT taxa) in response to organic enrichment.

7.5.8 An alternative to the ratio of EPT and Chironomidae abundance metric is the Indicator Assemblage Index (IAI) developed by Shackleford (1988). The IAI integrates the relative abundances of the EPT taxonomic groups and the relative abundances of chironomids and annelids upstream and downstream of a pollution source to evaluate impairment. The IAI may be a valuable metric in areas where the annelid community may fluctuate substantially in response to pollutant stress.

7.5.9 Percent Contribution of Dominant Taxon to the total number of organisms is an indication of community balance at the lowest possible taxonomic level. (The lowest positive taxonomic level is assumed to be genus or species in most instances). A community dominated by relatively few species would indicate environmental stress. Shackleford (1988) has modified this metric to reflect "dominants in common" (DIC) utilizing the dominant five taxa at the stations of comparison. The DIC will provide a measure of replacement or substitution between the reference community and the downstream station.

7.5.10 Community Similarity Indices are used in situations where reference communities exist. The reference community can be derived through sampling an upstream station or prediction for a region using a reference data base. Data sources or ecological data files may be available to establish a reference community for comparison. Several of the many similarity indices available are discussed below:

7.5.10.1 Community Loss Index measures the loss of benthic species between a reference station and the station of comparison. The community loss index was developed by Courtemanch and Davies (1987) and is an index of dissimilarity with values increasing as the degree of dissimilarity from the reference station increases. Values range from zero (0) to "infinity." Based on preliminary data analysis, this index provides greater discrimination than the following two community similarity indices. The formula for determining community loss index is:

$$I = \frac{a - c}{b}$$

where I = Coefficient of Community Loss, "a" is the number of taxa at the unimpacted site, "b" is the number of taxa at the study site, and "c" is the taxa common to "a" and "b". The result is a ratio of the number of taxa assumed lost due to the pollution source (a-c) to the number of taxa remaining including any new taxa.

7.5.10.2 Jaccard Coefficient of Community measures the degree of similarity in taxonomic composition between two stations in terms of taxa presence or absence and discriminates between highly similar collections (Jaccard, 1912). Coefficient values, ranging from 0 to 1.0, increase as the degree of similarity with the reference station increases. See Boesch (1977), and USEPA (1983) for more detail. The formula for the Jaccard Coefficient is:

$$\text{Jaccard Coefficient} = \frac{a}{a + b + c}$$

where

a = number of species common to both samples
b = number of species present in Sample B but not A
c = number of species present in Sample A but not B

Sample A = reference station
Sample B = station of comparison

7.5.10.3 The Index of Similarity (S) Between Two Samples has been used to determine whether shifts in community assemblages have occurred along a stream gradient or above and below a pollutional impact. The Index of Similarity can also be used as a quality assurance tool when evaluating variance in community assemblages between two control or reference sites. The inverse of the Index of Similarity is known as the Index of Dissimilarity. Both are reported as percentages and the formula is (Odum, 1971):

$$S = \frac{2C}{A + B}$$

Where A = Number of Species in Sample 1
B = Number of Species in Sample 2
C = Number of Species Common to both Species
1 - S = Index of Dissimilarity

7.5.10.4 The Pinkham and Pearson Community Similarity Index measures the degree of similarity in taxonomic composition in terms of taxa abundances and can be calculated with either percentages or numbers. A weighting factor can be added that assigns more significance to dominant species. See Pinkham and Pearson (1976) and USEPA (1983) for more detail. The formula is:

$$S.I._{ab} = \Sigma \frac{\text{Min}(X_{ia}, X_{ib})}{\text{Max}(X_{ia}, X_{ib})} \left[\frac{\frac{X_{ia}}{X_a} \cdot \frac{X_{ib}}{X_b} / 2}{\text{Weighting factor}} \right]$$

where X_{ia}, X_{ib} = number of individuals in the i^{th} species in sample A or B.

7.5.10.5 A Percent Similarity Method described by Gauch and Whittaker (1972) matches the benthic community structure of the site under study with an unimpacted site (control). It is a calculation of the degree to which the distribution of individuals within specific taxa in one site is similar to the distribution in another matched site. The value may range from zero (0) for sites with no taxa in common, to one (1) for identical communities.

$$P.S. = \frac{2\ \Sigma \quad \text{min.}\ (P_{ij}, P_{ik})}{(P_{ij} + P_{ik})}$$

where P.S. = Percent similarity, P_{ij} = Percentage of taxa "i" in community "j", and P_{ik} = Percentage of organisms of taxa "i" in community "k".

7.5.10.6 Other Community Similarity Indices include Spearman's Rank Correlation (Snedecor and Cochran, 1980); Moriseta's Index (Moriseta, 1959); Biotic Condition Index (Winget and Mangum, 1979); and Bray-Curtis Index (Bray and Curtis, 1957; Whittaker, 1952). Calculation of a chi-square "goodness of fit" (Cochran, 1952) may also be appropriate.

7.5.11 Presence and/or Absence of Specific Indicator Organisms is usually based upon a classification of organisms as either pollution sensitive (intolerant), facultative (variable), or tolerant (see paragraph 7.2.5). For example, usually stoneflies, mayflies, and caddisflies are considered sensitive or facultative and, therefore, are usually the first to suffer in a polluted environment. Sludgeworms and bloodworms, on the other hand, can tolerate very heavy pollutional loads.

7.5.11.1 The method differs from the biotic index of Hilsenhoff (1977, 1987) in that only selected indicator species are used to make decisions, whereas his biotic index used all the organisms in the samples.

7.5.11.2 A classic example of a system using the presence/absence criteria, is the Saprobien system (Kolkwitz and Marsson, 1908) which recognizes three basic zones of pollution ranging from a zone of heavy pollution (polysaprobic) characterized by a lack of dissolved oxygen, an abundance of bacteria, and the presence of a few tolerant species, to a zone of recovery (oligosaprobic) characterized by relatively pure water with a somewhat stable species diversity and dissolved oxygen concentration. This system was developed for use in Europe. Its usefulness is limited to organic pollutants in slow moving streams and is not always applicable to rivers and streams of the United States. A modification of the method was used in studies of the Illinois River (Richardson, 1928) and of a stream in southern Ohio (Gaufin and Tarzwell, 1956). A further modification of this method in combination with the biotic index was recently used by Rabeni et al. (1985) in the study of a Maine river. The results appear to be encouraging for wide use in this country. This approach is highly subjective and would naturally vary from one stream to another. It is also restricted to organic-type wastes.

7.5.12 Mean Number of Individuals per Sample is a simple means of comparing biological data. All of the individuals in all the replicate samples from one station are counted and divided by the number of replicates to yield the number of individuals per sample.

7.6 Statistical Methods

7.6.1 Graphical Examination of Data

Often the most elementary techniques are of the greatest use in data interpretation. Visual examination of data can point the way for more discriminatory analyses, or on the other hand, interpretations may become so obvious that further analysis is superfluous. In either case, graphical examination of data is often the most effortless way to obtain an initial examination of data and affords the chance to organize the data. Therefore, it is often done as a first step. Some commonly used techniques are presented below.

7.6.1.1 Raw Data

It is of utmost importance that raw data be recorded in a careful, logical, interpretable manner together with appropriate, but not superfluous, annotations. Note that although some annotations may be considered superfluous to the immediate intent of the data, they may not be so for other purposes. Any note that might aid in determining whether the data are comparable to other similar data, etc., should be recorded if possible.

7.6.1.2 Frequency Histograms

To construct a frequency histogram (see Freund, 1986) from the data, examine the raw data to determine the range, then establish intervals. Choose the intervals with care so they will be optimally integrative and differentiable. If the intervals are too wide, too many observations will be integrated into one interval and the picture will be hidden; if too narrow, too few will fall into one interval and a confusing overdifferentiation or overspreading of the data will result. It is often enlightening if the same data are plotted with the use of several interval sizes. Construct the intervals so that no doubt exists as to which interval an observation belongs, i.e., the end of one interval must not be the same number as the beginning of the next.

Although a frequency table contains all the information that a comparable histogram contains, the graphical value of a histogram is usually worth the small effort required for its construction. Histograms are more immediately interpretable. The height of each bar is the frequency of the interval; the width is the interval width.

7.6.1.3 Frequency Polygon

Another way to present essentially the same information as that in a frequency histogram is the use of a frequency polygon. Plot points at the height of the frequency and at the midpoint of the interval, and connect the points with straight lines.

7.6.1.4 Cumulative Frequency

Cumulative frequency plots are often useful in data interpretation. The height of a bar (frequency) is the sum of all frequencies up to and including the one being plotted. Thus, the first bar will be the same as the frequency histogram, the second bar equals the sum of the first and second bars of the frequency histogram, etc., and the last bar is the sum of all frequencies.

Closely related to the cumulative frequency histogram is the cumulative frequency distribution graph, a graph of relative frequencies. To obtain the cumulative graph, merely change the scale of the frequency axis on the cumulative frequency histogram. The scale change is made by dividing all values on the scale by the highest value on the scale.

The value of the cumulative frequency distribution graph is to allow relative frequency to be read, i.e., the fraction of observations less than or equal to some chosen value. Exercise caution in extrapolating from a cumulative frequency distribution to other situations. Always bear in mind that in spite of a planned lack of bias, each sample, or restricted set of samples, is subject to influences not accounted for and is therefore unique. This caution is all the more pertinent for cumulative frequency plots because they tend to smooth out some of the variation noticed in the frequency histogram. In addition, the phrase "fraction of observations less than or equal to some chosen value" can easily be read "fraction of time the observation is less than or equal to some chosen value." It is tempting to generalize from this reading and extend these results beyond their range of applicability.

7.6.1.5 Two-dimensional Graphs

Often data are taken where the observations are recorded as a pair (biomass and nutrient concentration). Here a quick plot of the set of pairs will usually be of value. The peaks and troughs, their frequency, together with intimate knowledge of the conditions of the study, might suggest something of biological interest, further statistical analysis, or further field or laboratory work.

7.6.1.6 In summary, carefully prepared tables and graphs may be important and informative steps in data analysis. The added effort is usually small, whereas gains in interpretive insight may be large. Therefore, graphic examination of data is a recommended procedure in the course of most investigations.

7.6.2 Sample Mean and Variance

7.6.2.1 Notation

Knowledge of certain computations and computational notations is essential to the use of statistical techniques. Some of the more basic of these will be briefly reviewed here.

To illustrate the computations, let us assume we have a set of data, i.e., a list of numeric values written down. Each of these values can be labeled by a set of numerals beginning with 1. Thus, the *first* of these values can be called X_1, the *second* X_2, etc., and the *last* one we call X_n,. The data values are labeled with consecutive numbers (recall from the definitions that these numeric values are observations), and there are n values in the set of data. A typical observation is X_i, where i may take any value between 1 and n, inclusive, and the subscript indicates which X is being referenced.

The sum of the numbers in a data set, such as our sample, is indicated in statistical computations by capital sigma, Σ. Associated with Σ are an operand (here, X_i), a subscript (here, i = 1), and a superscript (here, n).

$$\sum_{i=1}^{n} X_i$$

The subscript i= 1 indicates that the value of the operand X is to be the number labeled X_i in our data set and that this is to be the first observation of the sum. The superscript n indicates that the last number of the summation is to be the value of X_n the last X in our data set.

7.6.2.2 Calculation of the Sample Mean and Variance

Computations for the mean, variance, standard deviation, variance of the mean, and standard deviation of the mean (standard error) are presented below. Note that these are computations for a sample of n observations, i.e., they are statistics.

Note: The X_i's are squared, then the summation is performed in the first term

$$\textit{Mean}\ (\overline{X}):\quad \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\textit{Variance}\ (s^2):\quad s^2 = \frac{\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n}}{n-1}$$

of the numerator; in the second term, the sum of the X_i's is first formed, then the sum is squared, as indicated by the parenthesis.

$$\textit{Standard deviation}\ (s):\quad s = \sqrt{s^2}$$

$$\textit{Variance of the mean}\ (s_{\overline{X}}^2):\quad s_{\overline{X}}^2 = \frac{s^2}{n}$$

$$\textit{Standard deviation of the mean}\ (s_{\overline{X}}):\quad s_{\overline{X}} = \sqrt{s_{\overline{X}}^2} = \frac{s}{\sqrt{n}}$$

7.6.3 Rounding

The questions of rounding and the number of digits to carry through the calculations always arise in making statistical computations. Measurement data are approximations, since they are rounded when the measurements were taken; count data and binomial data are not subject to this type of approximation.

Observe the following rules when working with measurement or continuous data.

* When rounding numbers to some number of decimal places, first look at the digit to the right of the last place to be retained. If this number is greater than 5, the last place to be retained is rounded up by 1; if it is less than 5, do not change the last place - merely drop the extra places. To round to 2 decimal places:

Unrounded	Rounded
1.239	1.24
28.5849	28.58

* If the digit to the right of the last place to be retained is 5, then look at the second digit to the right of the last place to be kept, provided that the unrounded number is recorded with that digit as a significant digit. If the second digit to the right is greater than 0, then round the number up by 1 in the last place to be kept; if the second digit is 0, then look at the third digit, etc. To round to 1 place:

Unrounded	Rounded
13.251	13.3
13.25001	13.3

* If the number is recorded to only one place to the right of the last place to be kept, a special rule (odd-even rule) is followed to ensure that upward rounding occurs as frequently as downward rounding. The rule is: if the digit to the right of the last place to be kept is 5, and is the last digit of significance, round up when the last digit to be retained is *odd* and drop the 5 when the last digit to be retained is even. To round to 1 place:

Unrounded	Rounded
13.25	13.2
13.3500	13.4

Caution: all rounding must be made in 1 step to avoid introducing bias. For example the number 5.451 rounded to a whole number is clearly 5, but if the rounding were done in two steps it would first be rounded to 5.5 then 6.

Retention of significant figures in statistical computations can be summarized in three rules:

* Never use more significance for a raw data value than is warranted.

* During intermediate computations keep all significant figures for each data value, and carry the computations out in full.

* Round the final result to the accuracy set by the least accurate data value.

7.6.4 Tests of Hypotheses

7.6.4.1 Introduction

Often in biological field studies some aspect of the study is directed to answering a hypothetical question about a population (Allan, 1984). If the hypothesis is quantifiable, such as: "At the time of sampling, the standing crop of macroinvertebrates per basket at station 1 was the same as at station 2", then the hypothesis can be tested statistically. The question of drawing a sample in such a way that there is freedom from bias, so that such a test may be made, was discussed in Section 4, Selection of Sampling Stations.

There are many different types of hypothesis tests. Two basic categories of hypothesis tests are parametric tests, those based on the data following a specific distribution, and nonparametric tests, those based on relative rankings of the data. Three standard parametric tests of hypotheses will be presented here: the t-test, the χ^2 test, and the F-test. For information concerning nonparametric tests see Conover, 1980.

7.6.4.2 T-test

The t-test is used to compare a sample statistic (such as the mean) with some value for the purpose of making a judgment about the population as indicated by the sample. The comparison value may be the mean of another sample (in which case we are using the two samples to judge whether the two populations are the same). The form of the t-statistic is

$$t = \frac{\phi - \theta}{s_\phi}$$

where ϕ = some sample statistic; S_ϕ = the standard deviation of the sample statistic; and θ = the value to which the sample statistic is compared (the value of the null hypothesis).

The use of the t-test requires the use of t-tables. The t-table is a two-way table usually arranged with the column headings being the probability, α, of rejecting the null hypothesis when it is true, and the row headings being the degrees of freedom. Entry of the table at the correct probability level requires a discussion of two types of hypotheses testable using the t-statistic.

The null hypothesis is a hypothesis of no difference between a population parameter and another value. Suppose the hypothesis to be tested is that the mean, μ, of some population equals 10. Then we would write the null hypothesis (symbolized H_o) as

$$H_o : \mu = 10$$

Here 10 is the value of θ in the general form for the t-statistic. An alternative to the null hypothesis is now required. The investigator, viewing the experimental situation, determines the way in which this is stated. If the investigator merely wants to answer whether the sample indicates that μ = 10 or not, then the alternate hypothesis, H_a, is

$$H_a: \mu \neq 10$$

If it is known, for example, that μ cannot be less than 10, the H_a is

$$H_a: \mu > 10$$

and by similar reasoning the other possible H_a is

$$H_a: \mu < 10$$

Hence, there are two types of alternate hypotheses: one where the alternative is simply that the null hypothesis is false H_a: $\mu \neq 10$; the other, that the null hypothesis is false and, in addition, that the population parameter lies to one side or the other of the hypothesized value [H_a : μ (> or <) 10]. In the case of H_a: $\mu \neq 10$, the test is called a two-tailed test; in the case of either of the second types of alternate hypotheses,, the t-test is called a one-tailed test.

To use a t-table, it must be determined whether the column headings (probability of a larger value, or percentage points, or other means of expressing α) are set for one-tailed or two-tailed tests. Some tables are presented with both headings, and the terms "sign ignored" and "sign considered" are used. "Sign ignored" implies a two-tailed test, and "sign considered" implies a one-tailed test. Where tables are given for one-tailed tests, the column for any probability (or percentage) is the column appropriate to twice the probability for a two-tailed test. Hence, if a column heading 0.025 and the table is for one tailed tests, use this same column for 0.05 in a two tailed test (double any one-tailed test heading to get the proper two-tailed test heading; or conversely, halve the two-tailed test heading to obtain proper headings for one-tailed tests).

Testing H_o : μ = M (the population mean equals some value M):

$$t = \frac{\overline{X} - M}{s_{\overline{X}}}$$

where $\overline{X}$ is given by the sample mean; M = the hypothesized population mean; and $s_{\overline{X}}$ is given by the standard deviation (standard error) of the mean. The t-table is entered at the chosen probability level (often 0.05) and n-1 degrees of freedom, where n is the number of observations in the sample.

When the computed t-statistic exceeds the tabular value there is said to be a 1-α probability that H_o is false.

Testing H_o: $\mu_1 = \mu_2$ (the mean of the population from which sample 1 was taken equals the mean of the population from which sample 2 was taken):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1} - s_{\bar{X}_2}}$$

where $\bar{X}_1$ and $\bar{X}_2$ are the means from sample 1 and sample 2 respectively and $s_{\bar{X}_1} - s_{\bar{X}_2}$ is the standard error for the difference $\bar{X}_1 - \bar{X}_2$ calculated as follows:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{(n_1 + n_2 - 2)} \cdot \left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}$$

where s_1^2 and s_2^2 are variances of samples one and two respectively, and n_1 and n_2 are the number of observations for each sample.

For all conditions to be met where the t-test is applicable, the sample should have been selected from a population distributed as a normal distribution. Even if the population is not distributed normally, however, as sample size increases, the t-test approaches to applicability. If it is suspected that the population deviates too drastically from the normal, exercise care in the use of the t-test. Another assumption of the t-test is that the variances of the two populations are equal. Both the normality assumption and the equal variance assumption should be formally tested prior to using the t-test.

7.6.4.3 Chi-Square Test (χ^2)

The chi-square test is useful for statistically testing a hypothesis. Like t, χ^2 values may be found in mathematical and statistical tables tabulated in a
two-way arrangement. Usually, the column headings are probabilities of obtaining a larger χ^2 value when H_o is true, and the row headings are degrees of freedom.

If the calculated χ^2 exceeds the tabular value, then the null hypothesis is rejected. The chi square test is often used with the assumption of approximate normality in the population.

Chi-Square appears in two forms that differ not only in appearance, but that provide formats for different applications.

One form is useful in tests regarding hypotheses about σ^2:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_o^2} \quad \text{where } H_o: \sigma^2 = \sigma_o^2$$

The other form:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

where O = an observed value, and E = an expected (hypothesized) value, is especially useful in sampling from binomial and multinomial distributions, i.e., where the data may be classified into two or more categories (k).

Consider first a binomial situation. Suppose the Stenonema mayflies (2 species) from three stream riffle stations are pooled and the hypothesis of an equal ratio of the two species is tested based on the hypothetical data in Table 9.

Table 9. POOLED STENONEMA DATA FROM THREE RIFFLE STATIONS

Stenonema sp. 1	Stenonema sp. 2	Total
892* (919**)	946* (919**)	1838

* Observed value.
** Expected or hypothesized value.

To compute the hypothesized values (919 above) it is necessary to have formulated a null hypothesis. In this case it was H_o:No. Sp. 1 = No. Sp. 2 = (0.5) (Total).

Expected values are always computed based upon the null hypothesis. The computation for χ^2 is

$$\chi^2 = \frac{(892-919)^2 + (946-919)^2}{919} = 1.59 \; n.s.^*$$

*n.s. = not significant at α = 0.05

There is one degree of freedom for this test. Since the computed χ^2 is not greater than the tabulated χ^2 (3.84) for α =0.05, the null hypothesis is not rejected. This test, of course, applies equally well to data that has not been pooled, i.e., where the values are from two unpooled categories.

The information contained in each of the collections is partially obliterated by pooling. If the identity of the collections is maintained, two types of tests may be made; a test of the null hypothesis for each collection separately; and a test of interaction, i.e., whether the ratio depends upon the riffle from which the sample was obtained (Table 10).

With the use of the same null hypothesis, the following results are obtained. All tests were made at the α =0.01 level of significance. (Note: A significance level of 0.01 is used, instead of 0.05, to allow for the fact that multiple tests are being made within one experiment)

The individual χ^2's were computed, using the second form of chi square above, in separate tests of the hypothesis for each riffle. Note that the first two are not significant whereas the third is significant. This points to probable ecological differences among riffles, a possibility that would not have been discerned by pooling the data.

Table 10. STENONEMA DATA FROM THREE RIFFLE STATIONS

Riffle	Sp. 1	Sp. 2	Total	χ^2
1	346* (354)+	362 (354)	708	0.36 n.s.
2	302 (288)	274 (288)	576	1.30 n.s.
3	244 (277)	310 (277)	554	7.88
Total	892 (919)	946 (919)	1838	1.59 n.s.

* Observed values.
\+ Expected, or hypothesized values.

The test for interaction (dependence) is made by summing the individual χ^2's and subtracting the χ^2 obtained using totals, i.e.,

$$\chi^2 \text{ (interactions)} = \Sigma\, \chi^2 \text{ (individuals)} - \chi^2 \text{ (total)}$$

$$= 0.36 + 1.30 + 7.88 - 1.59$$

$$= 7.95$$

The degrees of freedom for the interaction χ^2 are the number of individual χ^2's minus one; in this case, two. This interaction χ^2 is significant, which indicates that the dominant species is indeed dependent upon the riffle.

Another χ^2 test may be illustrated by the following example. Suppose that comparable techniques were used to collect from four streams. With the use of three species common to all streams, it is desired to test the hypothesis that the three species occur in the same ratio regardless of stream, i.e., that their ratio is independent of stream (Table 11).

TABLE 11. OCCURRENCE OF THREE SPECIES OF MIDGES

Stream	Number of organisms			Frequency
	Species 1	Species 2	Species 3	
1	24* (21.7)+	12 (12.5)	30 (31.7)	66
2	15 (18.5)	14 (10.6)	27 (26.9)	56
3	28 (27.4)	15 (15.7)	40 (39.9)	83
4	20 (19.4)	9 (11.2)	30 (28.4)	59
Total	87	50	127	264
Expected ratio	87/264	50/264	127/264	

* Observed values.
\+ Expected or hypothesized.

To discuss the table above, O_{ij} = the observation for the i^{th} stream and the j^{th} species. Hence, O_{23} is the observation for stream two and species three. A similar indexing scheme applies to the expected values, E_{ij}. For the totals, a subscript replaced by a dot $E_{1.}$ symbolizes that summation has occurred for the observations indicated by that subscript. Hence, $O_{.2}$ is the total for species two (50); $O_{3.}$ is the total for stream three (83); and $O_{..}$ is the grand total (264).

Computations of expected values make use of the null hypothesis that the ratios are the same regardless of stream. The best estimate of this ratio for any species is $O_{.j}/O_{..}$, the ratio of the sum for species j to the total of all species. This ratio multiplied by the total for stream i gives the expected number of organisms of species j in stream i:

$$E_{ij} = \frac{O_{\cdot j}}{O_{\cdot\cdot}} \cdot (O_{i\cdot})$$

For example,

$$E_{12} = \frac{O_{\cdot 2}}{O_{\cdot\cdot}} \cdot (O_{1\cdot}) = \frac{50}{264} \cdot (66) = 12.5$$

χ^2 is computed as

$$\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 2.69 \quad (n.s.)$$

For this type of hypothesis, there are (rows – 1) (columns – 1) degrees of freedom, in this case

$$(4-1)\ (3-1) = 6$$

In the example, since the computed χ^2 is not greater than the tabulated $\chi^2(12.59)$ for α=0.05 the null hypothesis cannot be rejected. Thus, there is no evidence that the ratios among the organisms are different for different streams.

Tests of two types of hypotheses by χ^2 have been illustrated. The first type of hypothesis was one where there was a theoretical ratio, i.e., the ratio of sp.1 to sp.2 is 1:1. The second type of hypothesis was one where equal ratios were hypothesized, but the values of the ratios themselves were computed from the data. To draw the proper inference, it is important to make a distinction between these two types of hypotheses.

7.6.4.4 Analysis of Variance

Another form of hypothesis testing is the analysis of variance (ANOVA). The ANOVA is a powerful and general technique applicable to data from virtually any experimental or field study. There are restrictions, however, in the use of the technique. Experimental errors are assumed to be normally (or approximately normally) distributed about a mean of zero and have a common variance; they are also assumed to be independent (i.e., there should be no correlations among responses that are unaccounted for by the identifiable factors of the study or by the model). The effects tested must be assumed to be linearly additive. In practice these assumptions are rarely completely fulfilled, but the analysis of variance can be used unless significant departures from normality, or correlations among adjacent observations, or other types of measurement bias are suspected. It would be prudent, however, to check with a statistician regarding any uncertainties about the applicability of the test before issuing final reports or publications. Two simple but potentially useful examples of the analysis of variance are presented to illustrate the use of this technique.

7.6.4.4.1 Randomized Design

The analysis of variance for completely randomized designs provides a technique often useful in field studies. This test is commonly used for data derived from highly-controlled laboratory or field experiments where treatments are applied randomly to all experimental units, and the interest lies in whether or not the treatments significantly affected the response of the experimental units. This case may be of use in water quality studies, but in these studies the treatments are the conditions found, or are classifications based upon ecological criteria. Here the desire is to detect any differences in some type of measurement that might exist in conjunction with the field situation or the classifications or criteria.

For example, suppose it is desired to test whether the biomass of organisms in drift nets in a stream varies due to sampling time. Data from such a study are presented in Table 12.

In testing with the analysis of variance, as with other methods, a null hypothesis should be formulated. In this case the null hypothesis could be:

H_o: There are no differences in the biomass of organisms that may be attributed to time of sampling.

TABLE 12. MACROINVERTEBRATE BIOMASS COLLECTED AT DIFFERENT TIMES OF DAY FROM THE LITTLE MIAMI RIVER AT MILFORD, OHIO

Sampling Time (Time)	Replicate number	Biomass (mg dry wt.)
9:00AM – 1:00PM	1	1678
	2	1211
	3	1644
	4	1137
1:00AM – 4:00PM	1	1604
	2	1639
	3	2077
	4	2581
4:00PM – 7:00PM	1	4276
	2	2400
	3	3183
	4	3451

In utilizing the analysis of variance, the test for whether there are differences across time is made by comparing two types of variances, most often called "mean squares" in this context. Two mean squares are computed: one based upon the means for times; and one that is free of the effect of the means. In our example, a mean square for times is computed with the use of the averages (or totals) from the sampling time. The magnitude of this mean square is affected both by differences among the means and by differences among nets of the same time. The mean square within time is computed that has no contribution due to time differences. If the null hypothesis is true, then differences among sampling time do not exist and, therefore, they make no contribution to the mean square for times. Thus, both mean squares (between times and within times) are estimates of the same variance, and with repeated sampling, they would be expected to average to the same value. If the null hypothesis (H_o) is true, the ratio of these values is expected to equal one. If H_o is not true, i.e., if there are real differences due to the effect of times, then the mean square between times is affected by these differences and is expected to be the larger. The ratio in the second case is expected to be greater than one. The ratio of these two variances forms an F-test.

The analysis of variance is presented in Table 13A.

TABLE 13A. Generalized ANOVA Table

Source	df	SS
Total	$N-1$ *	$\sum_i \sum_j X_{ij}^2 - C$
Between Times	$t-1$	$[(\sum_i X_{i.}^2)/r_i] - C$
Within Times	$\sum_i (r_i - 1)$	Total SS – Stream SS

*The symbols are defined as N=total number of observations (nets); t=number of sampling times; r_i=number of nets for sample time i; X_{ij}=an observation (biomass of net j at sampling time i); X_i=sum of the observations for sampling time i; and

$$C\text{=correction for mean} = \frac{(\sum_i \sum_j X_{ij})^2}{N}$$

TABLE 13B. Completed ANOVA Table Using Macroinvertebrate Biomass Data

Source	df	SS	MS	F
Total	11	10,381,723		
Between Times	2	7,717,020	3,858,510	13.03**
Within Times	9	2,664,703	296,078	

** Significant at the 0.05 probability level.

The computations are:

$$C = \frac{(5670+7901+13310)^2}{12} = 60{,}215{,}680$$

$$\sum_i \sum_j X_{ij}^2 = (1678)^2 + (1211)^2 + \ldots + (3451)^2 = 70{,}597{,}403$$

$$\textit{Total SS} = 70{,}597{,}403 - 60{,}215{,}680 = 10{,}381{,}723$$

$$\sum_i \frac{X_{i.}^2}{r_i} = \frac{(5670)^2}{4} - \frac{(7901)^2}{4} + \frac{(13310)^2}{4} = 67{,}932{,}700$$

$$\textit{Between Times SS} = 67{,}932{,}700 - 60{,}215{,}680 = 7{,}717{,}020$$

$$\begin{aligned}\textit{Within Times SS} &= \textit{Total SS} - \textit{Between Times SS} \\ &= 10{,}381{,}723 - 7{,}717{,}020 = 2{,}664{,}703\end{aligned}$$

The mean squares (MS column) are computed by dividing the sums of squares (SS column) by its corresponding degrees of freedom (df column). The F-test is performed by computing the ratio, (Between Times MS)/(Within Times MS), in this case:

$$\frac{3,858,510}{296,078} = 13.03$$

When the calculated F value (13.03) is compared with the F values in the table (tabular F values) where df = 2 for the numerator and df = 9 for the denominator, we find that the calculated F exceeds the value of the tabular F for probability greater than 0.95. Thus the conclusion is that there are significant differences in biomass due to time of sampling.

Note that this analysis presumes good biological procedure and obviously cannot discriminate differences in sampling time from differences arising, for example, from the net having been placed in riffles with different current velocity. In general, the form of any analysis of variance derives from a model describing an observation in the experiment. In the example, the model, although not stated explicitly, assumed only one factor affecting a biomass measurement -- sampling time. If the model had included other factors, a more complicated analysis of variance would have resulted.

7.6.4.4.2 Factorial Design

Another application of a simple analysis of variance may be made where the factors are arranged factorially. Suppose a field study was conducted where the effect of a suspected toxic effluent upon the macroinvertebrate fauna of a river above and below a sewage treatment plant (STP) was in question (Tables 14A and 14B). Five samples were taken about one-quarter mile upstream and five one-quarter mile downstream in the spring, and the sampling scheme was repeated again in the summer. Standard statistical terminology refers to each of the combinations P_1T_1, P_2T_1, P_1T_2, and P_2T_2 as treatments or treatment combinations.

In planning for this field study, a null and alternate hypothesis should have been formed. In fact, whether stated explicitly or not, the null hypothesis was:

H_o: The toxic effluent has no effect upon the macroinvertebrate biomass collected.

This hypothesis is not stated in statistical terms and, therefore, only implicitly tells us what test to make. Let us look further at the analysis before attempting to state a null hypothesis in statistical terms.

In this study two factors are identifiable: times and positions. A study could have been done on each of the two factors separately, i.e., an attempt could have been made to distinguish whether there was a difference associated with times, assuming all other factors insignificant, and likewise with the positions. The example, used here, however, includes both factors simultaneously. Data are given for times and for positions but with the complication that we cannot assume that one is insignificant when studying the other. For the purpose of this study, whether there is a significant difference with times or on the other hand with positions, are questions that are of little interest. Of interest to this study is whether the above-below the STP difference varies with times. This type of contrast is termed a positions-times interaction. Thus, our null hypothesis is, in statistical terminology:

H_o: There is no significant interaction effect

An analysis of variance may be used to test this hypothesis. In order to meet the normality and homogeneity of variance assumptions of the analysis, the raw data were $\log_{10}$ transformed (Table 14B). All calculations are on the transformed data.

TABLE 14A. MACROINVERTEBRATE BIOMASS (GRAMS WET WT.)

Time Collected	Collected above STP	Collected below STP
Spring	437	193
	343	86
	337	119
	635	505
	373	171
Summer	888	28
	1778	18
	4332	117
	1078	26
	859	78

TABLE 14B. LOG_{10} TRANSFORMED DATA

Time Collected	Collected above STP	Collected below STP
Spring	2.64	2.28
	2.54	1.93
	2.53	2.08
	2.80	2.70
	2.57	2.23
Summer	2.95	1.45
	3.25	1.26
	3.64	2.07
	3.03	1.41
	2.93	1.89

TABLE 15. TREATMENT TOTALS FOR THE DATA OF TABLE 14B

Total	Positions		Times totals
	Above	Below	
Spring	13.08	11.22	24.3
Summer	15.8	8.08	23.88
	Positions		Grand
totals	28.88	19.3	48.18

Symbolically, an observation must have three indices specified to be completely identified: position, time, and sample number. Thus there are three subscripts: X_{ijk} is an observation at position i, time j, and from sample k. A value of 1 for i is above the STP; 2, below the STP; 1 for j is spring; 2, summer. A particular example is X_{123}, the third sample above the STP for the summer, or 3.64. A total (Table 15) is specified by using the dot notation. For the value of $X_{ij\cdot}$, then the individually sampled values for position i, time j are totaled. It is a total for a treatment combination. For example, the value of $X_{11\cdot}$, is 13.08, and the value of $X_{1\cdot\cdot}$, where sampling and times are both totaled to give the total for above the STP is 28.88. Treatment totals are presented in Table 15.

For a slight advantage in generality, let the following additional symbols apply: t = number of times of sampling (in this case t = 2); p = number of positions sample (in this case p = 2); s = number of samples per treatment combination; and n = the total number of observations.

The computations are:

Correction for the mean (CT):

$$CT = \frac{(\sum_i \sum_j \sum_k X_{ijk})^2}{n} = \frac{(48.18)^2}{20} = 116.06$$

$$TSS = \sum_i \sum_j \sum_k X_{ijk}^2 - CT = (2.64)^2 + (2.54)^2 + \cdots + (1.89)^2 - 116.06 = 7.54$$

Note that the divisor (5) may be factored out here, if desired, but where a different number of samples is taken for each treatment combination it should be left as above.

Position Sum of Squares (SSP):

$$SSP = \frac{(\sum_i X_{i..})^2}{st} - CT = \frac{(28.88)^2}{10} + \frac{(19.3)^2}{10} - 116.06 = 4.59$$

Times Sum of Squares (SST):

$$SST = \frac{(\sum_j X_{.j.})^2}{sp} - CT = \frac{(24.3)^2}{10} + \frac{(23.88)^2}{10} - 116.06 = 0.01$$

Interaction of Positions and Times of Sums Squares (SSPT):

$$SSPT = \frac{(\sum_i \sum_j X_{ij.})^2}{s} - SPS - SST - CT$$

$$\frac{(13.08)^2}{5} + \frac{(11.22)^2}{5} + \frac{(15.80)^2}{5} + \frac{(8.08)^2}{5} - 4.59 - 0.01 - 116.06 = 1.72$$

Error Sums of Squares (SSE):

$$SSE = TSS - SSP - SST - SSPT = 7.54 - 4.59 - 0.01 - 1.72 = 1.22$$

The completed ANOVA, including F tests, is given in Table 16. Although not important to this example, the main effects, positions and times, are tested for significance. The F table is entered with df = 1 for effect tested, and df = 16 for error. The positions effect is significant and the times effect is not significant, both tested at α=0.05. The interaction effect is significant, and we, therefore, conclude that there is a significant effect of the effluent changes across time on biomass.

TABLE 16. ANALYSIS OF VARIANCE TABLE FOR FIELD STUDY DATA OF TABLE 14

Source	df	SS	MS	F
Positions	1	4.59	4.59	57.38 **
Times	1	0.01	0.01	0.125
Positions X Times	1	1.72	1.72	21.51 **
Error	16	1.22	0.08	
Total	19	7.54		

** Significant at the 0.05 probability level.

7.6.5 Confidence Interval for Means

When means are computed in field studies, the desire often is to report them as intervals rather than as fixed numbers. This is entirely reasonable because computed means are virtually always derived from samples and are subject to the same uncertainty that is associated with the sample.

The correct computation of confidence intervals requires that the distribution of the observations be known. But very often approximations are close enough to correctness to be of use, and often are, or may be made to be, conservative. For computation of confidence intervals for the mean, the normal distribution is usually assumed to apply for several reasons: the central limit theorem assures us that with large samples the mean is likely to be approximately normally distributed; the required computations are well known and are easily applied; and when the normal distribution is known not to apply, suitable transformation of the data often is available to allow a valid application.

The confidence interval for a mean is an interval within which the true mean is said to have some stated probability of being found. If the probability of the mean not being in the interval is α (α could equal 0.1, 0.05, 0.01, or any probability value), then the statement may be written:

$$P\,(CL_1 < \mu < CL_2) = 1 - \alpha$$

This is read, "The probability that the lower confidence limit (CL_1) is less than the true mean (μ) and that the upper confidence limit (CL_2) is greater than the true mean, equals $1-\alpha$." However, we never know whether or not the true mean is actually included in the interval. So the confidence interval statement is really a statement about our procedure rather than about μ. It says that if we follow the procedure for repeated experiments, a proportion of those experiments equal to α will, by chance alone, fail to include the true mean between our limits. For example, if $\alpha=0.05$, we can expect 5 of 100 confidence intervals to fail to include the true mean.

To compute the limits, the sample mean, $\bar{X}$,; the standard error, s_x; and the degrees of freedom, n-1; must be known. A $t_{\alpha/2,n-1}$ value from tables of Student's t is obtained corresponding to n-1 degrees of freedom and probability α. The computation is:

$$CL_1 = \overline{X} - (t_{\alpha/2}) \cdot (s_{\overline{X}})$$
$$CL_2 = \overline{X} + (t_{\alpha/2}) \cdot (s_{\overline{X}})$$

7.6.6 Validating Normality and Homogeneity of Variance Assumptions[1]

7.6.6.1 Introduction

The t-test and the analysis of variance are parametric procedures based on the assumptions that the observations within treatments are independent and normally distributed, and that the variance of the observations is homogeneous across all groups of observations. These assumptions should be checked prior to using these tests, to determine if they have been met. Tests for validating the assumptions are provided in the following discussion. If the tests fail (if the data do not meet the assumptions), a non-parametric procedure such as Friedman's Test or Wilcoxon's Rank Sum Test may be more appropriate. However, the decision on whether to use parametric or non-parametric tests may be a judgment call, and a statistician should be consulted in selecting the analysis.

7.6.6.2 Test for Normal Distribution of Data

A formal test for normality is the Shapiro-Wilk's Test. The test statistic is obtained by dividing the square of an appropriate linear combination of the sample order statistics by the usual symmetric estimate of variance. The calculated W must be greater than zero and less than or equal to one. This test is recommended for a sample size of 50 or less. If the sample size is greater than 50, the Kolomogorov "D" statistic is recommended. An example of the Shapiro-Wilk's test is provided below.

The example uses macroinvertebrate biomass data. The same data are used

---

[1]Adapted and modified from USEPA, 1989

in the discussion of the homogeneity of variance determination and the one-way analysis of variance example. The data and the mean and standard deviation of the observations at each time are listed in Table 17.

The first step of the test for normality is to center the observations by subtracting the mean of all the observations within a concentration from each observation in that concentration. The centered observations are listed in Table 18.

Calculate the denominator, D, of the test statistic:

$$D = \sum_i X_i^2 - \frac{(\sum_i X_i)^2}{n}$$

$$D = 2{,}664{,}705 - \frac{(-3)^2}{12} = 2{,}664{,}704$$

Where: X_i = The i^{th} centered observations.

n = The total number of observations.

Order the centered observations from smallest to largest.

$$X^{(1)} - X^{(2)} \quad . \quad . \quad . \quad - \quad X^{(n)}$$

Where $X^{(i)}$ denotes the ith ordered observation. The ordered observations are listed in Table 19.

From Table 21, for the number of observations, n, obtain the coefficients a_1, a_2, ..., a_k, where k is approximately n/2. For the data in this example,

n = 12, k = 6. The a_i values are listed in Table 20.

Compute the test statistic, W, as follows:

$$W = \frac{1}{D}\left[\sum_{i=1}^{k} a_i \left(X^{(n-i+1)} - X^{(i)}\right)\right]$$

$$W = \frac{1}{2{,}664{,}704}(1610)^2 = 0.973$$

The differences, $X^{(n-i+1)} - X^{(i)}$, are listed in Table 20.

The decision rule for this test is to compare the critical value from Table 22 to the computed W. If the computed value is less than the critical value, conclude that the data are not normally distributed. For this example, the critical value at a significance level of 0.01 and 12 observations (n) is 0.805. The calculated value, 0.973, is not less than the critical value. Thus, the conclusion of the test is that the data are normally distributed.

TABLE 17. MACROINVERTEBRATE BIOMASS COLLECTED AT DIFFERENT TIMES OF DAY FROM THE LITTLE MIAMI RIVER AT MILFORD, OHIO

Sampling Time	Replicate number	Biomass (mg dry wt.)	S^2	$\bar{X}$
9:00AM - 1:00PM	1	1678	80,161	1418
	2	1211		
	3	1644		
	4	1137		
1:00AM - 4:00PM	1	1604	209,392	1975
	2	1639		
	3	2077		
	4	2581		
4:00PM - 7:00PM	1	4276	598,680	3328
	2	2400		
	3	3183		
	4	3451		

TABLE 18. EXAMPLE OF SHAPIRO-WILK'S TEST: CENTERED OBSERVATIONS

Sampling Time	Replicate			
	1	2	3	4
9:00AM - 1:00PM	260	-207	226	-281
1:00PM - 4:00PM	-371	-336	102	606
4:00PM - 7:00PM	948	-928	-145	123

TABLE 19. EXAMPLE OF SHAPIRO-WILK'S TEST: ORDERED OBSERVATIONS

i	$X_{(i)}$	i	$X_{(i)}$
1	−928	7	102
2	−371	8	123
3	−336	9	226
4	−280	10	260
5	−207	11	606
6	−145	12	948

TABLE 20. EXAMPLE OF SHAPIRO-WILK'S TEST: TABLE OF COEFFICIENTS AND DIFFERENCES

i	a_i	$X^{(n-i-1)} - X^{(i)}$	
1	.5475	1876	$X^{(12)}-X^{(1)}$
2	.3325	977	$X^{(11)}-X^{(2)}$
3	.2347	596	$X^{(10)}-X^{(3)}$
4	.1586	507	$X^{(9)}-X^{(4)}$
5	.0922	330	$X^{(8)}-X^{(5)}$
6	.0303	247	$X^{(7)}-X^{(6)}$

TABLE 21. COEFFICIENTS FOR THE SHAPIRO-WILKS TEST[1]

i \ n	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2	—	0.0000	0.1667	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291
3	—	—	—	0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4	—	—	—	—	—	0.0000	0.0561	0.0947	0.1224
5	—	—	—	—	—	—	—	0.0000	0.0399

i \ n	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	—	—	0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8	—	—	—	—	0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9	—	—	—	—	—	—	0.0000	0.0163	0.0303	0.0422
10	—	—	—	—	—	—	—	—	0.0000	0.0140

i \ n	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043	0.3018	0.2992	0.2968	0.2944
3	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533	0.2522	0.2510	0.2499	0.2487
4	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151	0.2152	0.2151	0.2150	0.2148
5	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836	0.1848	0.1857	0.1864	0.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	0.1092	0.1150	0.1201	0.1245	0.1283	0.1316	0.1346	0.1372	0.1395	0.1415
8	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089	0.1128	0.1162	0.1192	0.1219
9	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876	0.0923	0.0965	0.1002	0.1036
10	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672	0.0728	0.0778	0.0822	0.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	—	—	0.0000	0.0107	0.0200	0.0284	0.0358	0.0424	0.0483	0.0537
13	—	—	—	—	0.0000	0.0094	0.0178	0.0253	0.0320	0.0381
14	—	—	—	—	—	—	0.0000	0.0084	0.0159	0.0227
15	—	—	—	—	—	—	—	—	0.0000	0.0076

[1]Taken from Conover, 1980.

TABLE 21. COEFFICIENT FOR THE SHAPIRO-WILKS TEST (Continued)

i \ n	31	32	33	34	35	36	37	38	39	40
1	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015	0.3989	0.3964
2	0.2921	0.2898	0.2876	0.2854	0.2834	0.2813	0.2794	0.2774	0.2755	0.2737
3	0.2475	0.2462	0.2451	0.2439	0.2427	0.2415	0.2403	0.2391	0.2380	0.2368
4	0.2145	0.2141	0.2137	0.2132	0.2127	0.2121	0.2116	0.2110	0.2104	0.2098
5	0.1874	0.1878	0.1880	0.1882	0.1883	0.1883	0.1883	0.1881	0.1880	0.1878
6	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686	0.1689	0.1691
7	0.1433	0.1449	0.1463	0.1475	0.1487	0.1496	0.1505	0.1513	0.1520	0.1526
8	0.1243	0.1265	0.1284	0.1301	0.1317	0.1331	0.1344	0.1356	0.1366	0.1376
9	0.1066	0.1093	0.1118	0.1140	0.1160	0.1179	0.1196	0.1211	0.1225	0.1237
10	0.0899	0.0931	0.0961	0.0988	0.1013	0.1036	0.1056	0.1075	0.1092	0.1108
11	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947	0.0967	0.0986
12	0.0585	0.0629	0.0669	0.0706	0.0739	0.0770	0.0798	0.0824	0.0848	0.0870
13	0.0435	0.0485	0.0530	0.0572	0.0610	0.0645	0.0677	0.0706	0.0733	0.0759
14	0.0289	0.0344	0.0395	0.0441	0.0484	0.0523	0.0559	0.0592	0.0622	0.0651
15	0.0144	0.0206	0.0262	0.0314	0.0361	0.0404	0.0444	0.0481	0.0515	0.0546
16	0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372	0.0409	0.0444
17	—	—	0.0000	0.0062	0.0119	0.0172	0.0220	0.0264	0.0305	0.0343
18	—	—	—	—	0.0000	0.0057	0.0110	0.0158	0.0203	0.0244
19	—	—	—	—	—	—	0.0000	0.0053	0.0101	0.0146
20	—	—	—	—	—	—	—	—	0.0000	0.0049

i \ n	41	42	43	44	45	46	47	48	49	50
1	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3770	0.3751
2	0.2719	0.2701	0.2684	0.2667	0.2651	0.2635	0.2620	0.2604	0.2589	0.2574
3	0.2357	0.2345	0.2334	0.2323	0.2313	0.2302	0.2291	0.2281	0.2271	0.2260
4	0.2091	0.2085	0.2078	0.2072	0.2065	0.2058	0.2052	0.2045	0.2038	0.2032
5	0.1876	0.1874	0.1871	0.1868	0.1865	0.1862	0.1859	0.1855	0.1851	0.1847
6	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
7	0.1531	0.1535	0.1539	0.1542	0.1545	0.1548	0.1550	0.1551	0.1553	0.1554
8	0.1384	0.1392	0.1398	0.1405	0.1410	0.1415	0.1420	0.1423	0.1427	0.1430
9	0.1249	0.1259	0.1269	0.1278	0.1286	0.1293	0.1300	0.1306	0.1312	0.1317
10	0.1123	0.1136	0.1149	0.1160	0.1170	0.1180	0.1189	0.1197	0.1205	0.1212
11	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
12	0.0891	0.0909	0.0927	0.0943	0.0959	0.0972	0.0986	0.0998	0.1010	0.1020
13	0.0782	0.0804	0.0824	0.0842	0.0860	0.0876	0.0892	0.0906	0.0919	0.0932
14	0.0677	0.0701	0.0724	0.0745	0.0765	0.0783	0.0801	0.0817	0.0832	0.0846
15	0.0575	0.0602	0.0628	0.0651	0.0673	0.0694	0.0713	0.0731	0.0748	0.0764
16	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0667	0.0685
17	0.0379	0.0411	0.0442	0.0471	0.0497	0.0522	0.0546	0.0568	0.0588	0.0608
18	0.0283	0.0318	0.0352	0.0383	0.0412	0.0439	0.0465	0.0489	0.0511	0.0532
19	0.0188	0.0227	0.0263	0.0296	0.0328	0.0357	0.0385	0.0411	0.0436	0.0459
20	0.0094	0.0136	0.0175	0.0211	0.0245	0.0277	0.0307	0.0335	0.0361	0.0386
21	0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
22		—	0.0000	0.0042	0.0081	0.0118	0.0153	0.0185	0.0215	0.0244
23	—	—	—	—	0.0000	0.0039	0.0076	0.0111	0.0143	0.0174
24	—	—	—	—	—	—	0.0000	0.0037	0.0071	0.0104
25	—	—	—	—	—	—	—	—	0.0000	0.0035

TABLE 22. QUANTILES OF THE SHAPIRO-WILKS TEST STATISTIC[1]

n	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
4	0.687	0.707	0.748	0.792	0.935	0.987	0.992	0.996	0.997
5	0.686	0.715	0.762	0.806	0.927	0.979	0.986	0.991	0.993
6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
7	0.730	0.760	0.803	0.838	0.928	0.972	0.979	0.985	0.988
8	0.749	0.778	0.818	0.851	0.932	0.972	0.978	0.984	0.987
9	0.764	0.791	0.829	0.859	0.935	0.972	0.978	0.984	0.986
10	0.781	0.806	0.842	0.869	0.938	0.972	0.978	0.983	0.986
11	0.792	0.817	0.850	0.876	0.940	0.973	0.979	0.984	0.986
12	0.805	0.828	0.859	0.883	0.943	0.973	0.979	0.984	0.986
13	0.814	0.837	0.866	0.889	0.945	0.974	0.979	0.984	0.986
14	0.825	0.846	0.874	0.895	0.947	0.975	0.980	0.984	0.986
15	0.835	0.855	0.881	0.901	0.950	0.975	0.980	0.984	0.987
16	0.844	0.863	0.887	0.906	0.952	0.976	0.981	0.985	0.987
17	0.851	0.869	0.892	0.910	0.954	0.977	0.981	0.985	0.987
18	0.858	0.874	0.897	0.914	0.956	0.978	0.982	0.986	0.988
19	0.863	0.879	0.901	0.917	0.957	0.978	0.982	0.986	0.988
20	0.868	0.884	0.905	0.920	0.959	0.979	0.983	0.986	0.988
21	0.873	0.888	0.908	0.923	0.960	0.980	0.983	0.987	0.989
22	0.878	0.892	0.911	0.926	0.961	0.980	0.984	0.987	0.989
23	0.881	0.895	0.914	0.928	0.962	0.981	0.984	0.987	0.989
24	0.884	0.898	0.916	0.930	0.963	0.981	0.984	0.987	0.989
25	0.888	0.901	0.918	0.931	0.964	0.981	0.985	0.988	0.989
26	0.891	0.904	0.920	0.933	0.965	0.982	0.985	0.988	0.989
27	0.894	0.906	0.923	0.935	0.965	0.982	0.985	0.988	0.990
28	0.896	0.908	0.924	0.936	0.966	0.982	0.985	0.988	0.990
29	0.898	0.910	0.926	0.937	0.966	0.982	0.985	0.988	0.990
30	0.900	0.912	0.927	0.939	0.967	0.983	0.985	0.988	0.990
31	0.902	0.914	0.929	0.940	0.967	0.983	0.986	0.988	0.990
32	0.904	0.915	0.930	0.941	0.968	0.983	0.986	0.988	0.990
33	0.906	0.917	0.931	0.942	0.968	0.983	0.986	0.989	0.990
34	0.908	0.919	0.933	0.943	0.969	0.983	0.986	0.989	0.990
35	0.910	0.920	0.934	0.944	0.969	0.984	0.986	0.989	0.990
36	0.912	0.922	0.935	0.945	0.970	0.984	0.986	0.989	0.990
37	0.914	0.924	0.936	0.946	0.970	0.984	0.987	0.989	0.990
38	0.916	0.925	0.938	0.947	0.971	0.984	0.987	0.989	0.990
39	0.917	0.927	0.939	0.948	0.971	0.984	0.987	0.989	0.991
40	0.919	0.928	0.940	0.949	0.972	0.985	0.987	0.989	0.991
41	0.920	0.929	0.941	0.950	0.972	0.985	0.987	0.989	0.991
42	0.922	0.930	0.942	0.951	0.972	0.985	0.987	0.989	0.991
43	0.923	0.932	0.943	0.951	0.973	0.985	0.987	0.990	0.991
44	0.924	0.933	0.944	0.952	0.973	0.985	0.987	0.990	0.991
45	0.926	0.934	0.945	0.953	0.973	0.985	0.988	0.990	0.991
46	0.927	0.935	0.945	0.953	0.974	0.985	0.988	0.990	0.991
47	0.928	0.936	0.946	0.954	0.974	0.985	0.988	0.990	0.991
48	0.929	0.937	0.947	0.954	0.974	0.985	0.988	0.990	0.991
49	0.929	0.937	0.947	0.955	0.974	0.985	0.988	0.990	0.991
50	0.930	0.938	0.947	0.955	0.974	0.985	0.988	0.990	0.991

[1]Taken from Conover, 1980.

7.6.6.3 Test for Homogeneity of Variance

For the analysis of variance, the variances of the data obtained for each group of observations are assumed to be equal. Bartlett's Test is a formal test of this assumption. In using this test, it is assumed that the data are normally distributed.

The data used in this example are biomass data from the one-way analysis of variance example and the Shapiro-Wilk's Test example. These data are listed in Table 17, together with the calculated sample variance for each group of observations.

The test statistic for Bartlett's Test (Snedecor and Cochran, 1980) is as follows:

$$B = \frac{\left[\left(\sum_{i=1}^{p} V_i\right)\ln \overline{S}^2 - \sum_{i=1}^{p}\left(V_i \ln S_i^2\right)\right]}{C}$$

Where: V_i = Degrees of freedom for each time
p = Number of levels of times

$\overline{S}^2$ = The average of the individual variances.
ln = Log_e

$$C = 1 + \left[\frac{1}{3(p-1)}\right]\left[\sum_{i=1}^{p}\frac{1}{V_i} - \frac{1}{\sum_{i=1}^{p} V_i}\right]$$

Since B is approximately distributed as chi-square with p – 1 degrees of freedom when the variances are equal, the appropriate critical value is obtained from a table of the chi-square distribution for p – 1 degrees of freedom and a significance level of α. If B is less than the critical value then the variances are assumed to be equal.

For the data in this example, $V_i = 4 - 1 = 3$, $p = 3$, $\overline{S}^2 = 296{,}078$, and $C = 1.148$. The calculated value is:

$$B = \frac{\left[\left(\sum_{i=1}^{3} 3\right)\ln \overline{S}^2 - 3\sum_{i=1}^{3}\left(\ln S_i^2\right)\right]}{1.148}$$

$$B = \frac{9(12.598) - 3(36.846)}{1.148} = 2.477$$

Since B is approximately distributed as chi-square with 2 degrees of freedom when the variances are equal, the appropriate critical value for the test is 9.210 (see a χ^2 table) for a significance level of 0.01. Since B = 2.477 is less than the critical value of 9.210, conclude that the variances are not different.

7.6.6.4 Transformations of the Data

When the assumptions of normality and/or homogeneity of variance are not met, transformations of the data may remedy the problem, so that the data can be analyzed by parametric procedures, rather than a non-parametric technique such as Friedman's Test or Wilcoxon's Rank Sum Test. Examples of transformations include log, square root, arc sine square root, and reciprocals. After the data have been transformed, Shapiro-Wilk's and Bartlett's test should be performed on the transformed observations to determine whether the assumptions of normality and/or homogeneity of variance are met.

Table 23 is reproduced here with permission from Lloyd, Zar, and Karr (1968) for use in calculating mean diversity ($\bar{d}$) (see 7.3.10, page 114). To use the table, find the number of individuals (n) in column 1 and read the log of that number in column 3 (n log n).

TABLE 23. FUNCTIONS FOR CALCULATING SPECIES DIVERSITY AND (FOR PERFECTLY RANDOM SAMPLING) ITS STANDARD ERROR LOGARITHMS ARE TO BASE 10. TABLE VALUES ARE ACCURATE TO WITHIN ±1 IN THE EIGHTH SIGNIFICANT FIGURE.

n	log n!	n log n	n log² n
1	.0000	.0000	.0000
2	.3010	.6021	.1812
3	.7782	1.4314	.6829
4	1.3802	2.4082	1.4499
5	2.0792	3.4949	2.4428
6	2.8573	4.6689	3.6331
7	3.7024	5.9157	4.9993
8	4.6055	7.2247	6.5246
9	5.5598	8.5882	8.1952
10	6.5598	10.0000	10.0000
11	7.6012	11.4553	11.9295
12	8.6803	12.9502	13.9756
13	9.7943	14.4813	16.1313
14	10.9404	16.0458	18.3905
15	12.1165	17.6414	20.7479
16	13.3206	19.2659	23.1985
17	14.5511	20.9176	25.7381
18	15.8063	22.5949	28.3628
19	17.0851	24.2963	31.0690
20	18.3861	26.0206	33.8536
21	19.7083	27.7666	36.7135
22	21.0508	29.5333	39.6462
23	22.4125	31.3197	42.6490
24	23.7927	33.1251	45.7196
25	25.1906	34.9485	48.8559
26	26.6056	36.7893	52.0559
27	28.0370	38.6468	55.3177
28	29.4841	40.5204	58.6395
29	30.9465	42.4095	62.0196
30	32.4237	44.3136	65.4566
31	33.9150	46.2322	68.9490
32	35.4202	48.1648	72.4952
33	36.9387	50.1110	76.0942
34	38.4702	52.0703	79.7445
35	40.0142	54.0424	83.4451
36	41.5705	56.0269	87.1948
37	43.1387	58.0235	90.9925
38	44.7185	60.0318	94.8372
39	46.3096	62.0515	98.7280
40	47.9116	64.0824	102.6638
41	49.5244	66.1241	106.6439
42	51.1477	68.1765	110.6674
43	52.7811	70.2391	114.7334
44	54.4246	72.3119	118.8412
45	56.0778	74.3946	122.9900
46	57.7406	76.4869	127.1791
47	59.4127	78.5886	131.4078
48	61.0939	80.6996	135.6755
49	62.7841	82.8196	139.9814
50	64.4831	84.9485	144.3250
51	66.1906	87.0861	148.7056
52	67.9066	89.2322	153.1227
53	69.6309	91.3866	157.5757
54	71.3633	93.5493	162.0642
55	73.1037	95.7199	166.5874
56	74.8519	97.8985	171.1450
57	76.6077	100.0849	175.7365
58	78.3712	102.2788	180.3613
59	80.1420	104.4803	185.0191
60	81.9202	106.6891	189.7093
61	83.7055	108.9051	194.4316
62	85.4979	111.1283	199.1854
63	87.2972	113.3585	203.9705
64	89.1034	115.5955	208.7863
65	90.9163	117.8394	213.6326
66	92.7359	120.0899	218.5088
67	94.5619	122.3470	223.4148
68	96.3945	124.6106	228.3500
69	98.2333	126.8806	233.3143
70	100.0784	129.1569	238.3071
71	101.9297	131.4393	243.3282
72	103.7870	133.7279	248.3772
73	105.6503	136.0226	253.4540
74	107.5196	138.3231	258.5580
75	109.3946	140.6296	263.6891
76	111.2754	142.9418	268.8469
77	113.1619	145.2598	274.0312
78	115.0540	147.5834	279.2417
79	116.9516	149.9125	284.4781
80	118.8547	152.2472	289.7401
81	120.7632	154.5873	295.0275
82	122.6770	156.9327	300.3400
83	124.5961	159.2835	305.6774
84	126.5204	161.6395	311.0395
85	128.4498	164.0006	316.4259
86	130.3843	166.3669	321.8364
87	132.3238	168.7382	327.2709
88	134.2683	171.1145	332.7291
89	136.2177	173.4957	338.2108
90	138.1719	175.8818	343.7157
91	140.1310	178.2728	349.2437
92	142.0948	180.6685	354.7946
93	144.0632	183.0689	360.3680
94	146.0364	185.4740	365.9640
95	148.0141	187.8837	371.5821
96	149.9964	190.2980	377.2223
97	151.9831	192.7169	382.8844
98	153.9744	195.1402	388.5682
99	155.9700	197.5679	394.2734
100	157.9700	200.0000	400.0000
101	159.9743	202.4365	405.7477
102	161.9829	204.8772	411.5164
103	163.9958	207.3222	417.3059
104	166.0128	209.7715	423.1160
105	168.0340	212.2249	428.9466
106	170.0593	214.6824	434.7976
107	172.0887	217.1441	440.6686
108	174.1221	219.6098	446.5597
109	176.1595	222.0795	452.4706
110	178.2009	224.5532	458.4013
111	180.2462	227.0309	464.3514
112	182.2955	229.5124	470.3210
113	184.3485	231.9979	476.3098
114	186.4054	234.4872	482.3178
115	188.4661	236.9803	488.3447
116	190.5306	239.4771	494.3905
117	192.5988	241.9777	500.4550
118	194.6707	244.4821	506.5380
119	196.7462	246.9901	512.6395
120	198.8254	249.5017	518.7594
121	200.9082	252.0170	524.8974
122	202.9945	254.5359	531.0535
123	205.0844	257.0583	537.2275
124	207.1779	259.5843	543.4194
125	209.2748	262.1138	549.6290
126	211.3751	264.6467	555.8561
127	213.4790	267.1831	562.1007
128	215.5862	269.7229	568.3627
129	217.6967	272.2661	574.6420
130	219.8107	274.8126	580.9383
131	221.9280	277.3625	587.2517
132	224.0485	279.9158	593.5821
133	226.1724	282.4723	599.9292
134	228.2995	285.0320	606.2930
135	230.4298	287.5951	612.6735
136	232.5634	290.1613	619.0704
137	234.7001	292.7307	625.4837
138	236.8400	295.3033	631.9134
139	238.9830	297.8791	638.3592
140	241.1291	300.4579	644.8212

TABLE 23. (Continued)

n	log n!	n log n	n log² n
141	243.2783	303.0399	651.2991
142	245.4306	305.6249	657.7930
143	247.5860	308.2131	664.3027
144	249.7443	310.8042	670.8281
145	251.9057	313.3984	677.3692
146	254.0700	315.9955	683.9258
147	256.2374	318.5956	690.4979
148	258.4076	321.1987	697.0853
149	260.5808	323.8048	703.6880
150	262.7569	326.4137	710.3060
151	264.9359	329.0255	716.9390
152	267.1177	331.6402	723.5871
153	269.3024	334.2578	730.2501
154	271.4899	336.8782	736.9280
155	273.6803	339.5014	743.6207
156	275.8734	342.1274	750.3281
157	278.0693	344.7562	757.0501
158	280.2679	347.3878	763.7867
159	282.4693	350.0221	770.5377
160	284.6735	352.6592	777.3032
161	286.8803	355.2990	784.0830
162	289.0898	357.9414	790.8770
163	291.3020	360.5866	797.6852
164	293.5168	363.2344	804.5075
165	295.7343	365.8849	811.3438
166	297.9544	368.5379	818.1941
167	300.1771	371.1936	825.0582
168	302.4024	373.8520	831.9362
169	304.6303	376.5129	838.8280
170	306.8608	379.1763	845.7334
171	309.0938	381.8423	852.6524
172	311.3293	384.5109	859.5850
173	313.5674	387.1820	866.5311
174	315.8079	389.8556	873.4906
175	318.0509	392.5317	880.4634
176	320.2965	395.2102	887.4496
177	322.5444	397.8913	894.4489
178	324.7948	400.5748	901.4615
179	327.0477	403.2607	908.4871
180	329.3030	405.9491	915.5257
181	331.5606	408.6398	922.5774
182	333.8207	411.3330	929.6419
183	336.0832	414.0285	936.7193
184	338.3480	416.7265	943.8096
185	340.6152	419.4268	950.9125
186	342.8847	422.1294	958.0282
187	345.1565	424.8344	965.1564
188	347.4307	427.5417	972.2973
189	349.7071	430.2513	979.4506
190	351.9859	432.9632	986.6164
191	354.2669	435.6774	993.7946
192	356.5502	438.3938	1000.9852
193	358.8358	441.1126	1008.1880
194	361.1236	443.8335	1015.4031
195	363.4136	446.5567	1022.6304
196	365.7059	449.2822	1029.8698
197	368.0003	452.0098	1037.1213
198	370.2970	454.7397	1044.3849
199	372.5939	457.4718	1051.6604
200	374.8969	460.2060	1058.9478
201	377.2001	462.9424	1066.2471
202	379.5054	465.6810	1073.5583
203	381.8129	468.4217	1080.8812
204	384.1226	471.1646	1088.2159
205	386.4343	473.9095	1095.5622
206	388.7482	476.6566	1102.9202
207	391.0642	479.4059	1110.2897
208	393.3822	482.1572	1117.6709
209	395.7024	484.9106	1125.0635
210	398.0246	487.6661	1132.4675
211	400.3489	490.4236	1139.8829
212	402.6752	493.1832	1147.3098
213	405.0036	495.9449	1154.7479
214	407.3340	498.7085	1162.1973
215	409.6664	501.4743	1169.6579
216	412.0009	504.2420	1177.1296
217	414.3373	507.0118	1184.6125
218	416.6758	509.7835	1192.1066
219	419.0162	512.5573	1199.6116
220	421.3587	515.3330	1207.1277
221	423.7031	518.1107	1214.6547
222	426.0494	520.8904	1222.1926
223	428.3977	523.6720	1229.7415
224	430.7480	526.4556	1237.3011
225	433.1002	529.2411	1244.8716
226	435.4543	532.0285	1252.4528
227	437.8103	534.8179	1260.0447
228	440.1682	537.6091	1267.6473
229	442.5281	540.4023	1275.2606
230	444.8898	543.1974	1282.8844
231	447.2534	545.9944	1290.5188
232	449.6189	548.7932	1298.1637
233	451.9862	551.5939	1305.8192
234	454.3555	554.3965	1313.4850
235	456.7265	557.2009	1321.1613
236	459.0994	560.0072	1328.8479
237	461.4742	562.8154	1336.5448
238	463.8508	565.6253	1344.2521
239	466.2292	568.4371	1351.9696
240	468.6094	571.2507	1359.6973
241	470.9914	574.0661	1367.4352
242	473.3752	576.8833	1375.1833
243	475.7608	579.7023	1382.9415
244	478.1482	582.5231	1390.7098
245	480.5374	585.3457	1398.4881
246	482.9283	588.1700	1406.2764
247	485.3210	590.9961	1414.0747
248	487.7154	593.8240	1421.8829
249	490.1116	596.6536	1429.7010
250	492.5096	599.4850	1437.5291
251	494.9093	602.3181	1445.3669
252	497.3107	605.1529	1453.2146
253	499.7138	607.9895	1461.0720
254	502.1186	610.8278	1468.9392
255	504.5252	613.6677	1476.8161
256	506.9334	616.5094	1484.7026
257	509.3433	619.3528	1492.5988
258	511.7549	622.1979	1500.5047
259	514.1682	625.0446	1508.4201
260	516.5832	627.8931	1516.3450
261	518.9999	630.7432	1524.2795
262	521.4182	633.5949	1532.2234
263	523.8381	636.4484	1540.1769
264	526.2597	639.3034	1548.1397
265	528.6830	642.1602	1556.1119
266	531.1078	645.0185	1564.0936
267	533.5344	647.8785	1572.0845
268	535.9625	650.7401	1580.0847
269	538.3922	653.6034	1588.0943
270	540.8236	656.4682	1596.1130
271	543.2566	659.3347	1604.1410
272	545.6912	662.2027	1612.1782
273	548.1273	665.0724	1620.2245
274	550.5651	667.9437	1628.2800
275	553.0044	670.8165	1636.3446
276	555.4453	673.6909	1644.4182
277	557.8878	676.5669	1652.5009
278	560.3318	679.4445	1660.5927
279	562.7774	682.3236	1668.6934
280	565.2246	685.2042	1676.8031
281	567.6733	688.0863	1684.9217
282	570.1235	690.9702	1693.0492
283	572.5753	693.8556	1701.1856
284	575.0287	696.7424	1709.3309
285	577.4835	699.6308	1717.4850
286	579.9399	702.5207	1725.6479
287	582.3977	705.4121	1733.8196
288	584.8571	708.3050	1742.0001
289	587.3180	711.1995	1750.1893
290	589.7804	714.0954	1758.3871
291	592.2443	716.9929	1766.5937
292	594.7097	719.8918	1774.8089
293	597.1766	722.7922	1783.0327
294	599.6449	725.6941	1791.2651
295	602.1147	728.5975	1799.5061
296	604.5860	731.5023	1807.7557
297	607.0588	734.4087	1816.0138
298	609.5330	737.3164	1824.2803
299	612.0087	740.2257	1832.5554
300	614.4858	743.1364	1840.8389
301	616.9644	746.0485	1849.1308
302	619.4444	748.9621	1857.4312
303	621.9258	751.8771	1865.7399
304	624.4087	754.7936	1874.0570
305	626.8930	757.7115	1882.3824
306	629.3787	760.6308	1890.7162
307	631.8659	763.5515	1899.0582
308	634.3544	766.4736	1907.4085
309	636.8444	769.3972	1915.7670
310	639.3357	772.3221	1924.1337
311	641.8285	775.2485	1932.5087
312	644.3226	778.1762	1940.8518
313	646.8182	781.1034	1949.2831
314	649.3151	784.0339	1957.6825
315	651.8134	786.9678	1966.0900
316	654.3131	789.9011	1974.5056
317	656.8142	792.8358	1982.9293
318	659.3166	795.7718	1991.3610
319	661.8204	798.7092	1999.8007
320	664.3255	801.6480	2008.2484
321	666.8320	804.5881	2016.7041
322	669.3399	807.5296	2025.1678
323	671.8491	810.4724	2033.6394
324	674.3596	813.4166	2042.1189
325	676.8715	816.3621	2050.6064
326	679.3847	819.3089	2059.1016
327	681.8993	822.2571	2067.6048
328	684.4152	825.2066	2076.1157
329	686.9324	828.1575	2084.6345
330	689.4509	831.1096	2093.1611
331	691.9707	834.0631	2101.6954
332	694.4918	837.0178	2110.2375
333	697.0143	839.9739	2118.7874
334	699.5380	842.9313	2127.3449
335	702.0631	845.8900	2135.9102
336	704.5894	848.8500	2144.4831
337	707.1170	851.8113	2153.0636
338	709.6460	854.7738	2161.6518
339	712.1762	857.7377	2170.2477
340	714.7076	860.7028	2178.8510
341	717.2404	863.6692	2187.4620
342	719.7744	866.6369	2196.0806
343	722.3097	869.6059	2204.7067
344	724.8463	872.5761	2213.3403
345	727.3841	875.5476	2221.9814
346	729.9232	878.5203	2230.6299
347	732.4635	881.4943	2239.2860
348	735.0051	884.4696	2247.9495
349	737.5479	887.4461	2256.6204
350	740.0920	890.4238	2265.2988
351	742.6373	893.4028	2273.9845
352	745.1838	896.3830	2282.6776
353	747.7316	899.3645	2291.3780
354	750.2806	902.3472	2300.0858
355	752.8308	905.3311	2308.8009
356	755.3823	908.3162	2317.5233
357	757.9349	911.3026	2326.2531
358	760.4888	914.2901	2334.9900
359	763.0439	917.2789	2343.7342
360	765.6002	920.2689	2352.4857
361	768.1577	923.2601	2361.2444
362	770.7164	926.2525	2370.0102
363	773.2764	929.2461	2378.7832
364	775.8375	932.2409	2387.5634
365	778.3997	935.2369	2396.3508
366	780.9632	938.2341	2405.1453
367	783.5279	941.2324	2413.9469
368	786.0937	944.2320	2422.7556

TABLE 23. (Continued)

n	log n!	n log n	n log² n
369	788.6608	947.2327	2431.5714
370	791.2290	950.2346	2440.3942
371	793.7983	953.2377	2449.2261
372	796.3689	956.2420	2458.0610
373	798.9406	959.2474	2466.9050
374	801.5135	962.2540	2475.7559
375	804.0875	965.2617	2484.6139
376	806.6627	968.2706	2493.4787
377	809.2390	971.2807	2502.3506
378	811.8165	974.2919	2511.2294
379	814.3952	977.3043	2520.1151
380	816.9749	980.3178	2529.0077
381	819.5559	983.3324	2537.9072
382	822.1379	986.3482	2546.8135
383	824.7211	989.3651	2555.7268
384	827.3055	992.3832	2564.6469
385	829.8909	995.4024	2573.5738
386	832.4775	998.4227	2582.5075
387	835.0652	1001.4441	2591.4480
388	837.6540	1004.4667	2600.3953
389	840.2440	1007.4904	2609.3493
390	842.8351	1010.5152	2618.3102
391	845.4272	1013.5411	2627.2777
392	848.0205	1016.5681	2636.2520
393	850.6149	1019.5963	2645.2330
394	853.2104	1022.6255	2654.2206
395	855.8070	1025.6558	2663.2150
396	858.4047	1028.6873	2672.2160
397	861.0035	1031.7198	2681.2237
398	863.6034	1034.7534	2690.2380
399	866.2044	1037.7882	2699.2589
400	868.8064	1040.8240	2708.2865
401	871.4096	1043.8609	2717.3206
402	874.0138	1046.8989	2726.3613
403	876.6191	1049.9379	2735.4086
404	879.2255	1052.9781	2744.4624
405	881.8329	1056.0193	2753.5228
406	884.4415	1059.0616	2762.5897
407	887.0510	1062.1049	2771.6631
408	889.6617	1065.1493	2780.7430
409	892.2734	1068.1948	2789.8293
410	894.8862	1071.2414	2798.9222
411	897.5001	1074.2890	2808.0215
412	900.1150	1077.3376	2817.1272
413	902.7309	1080.3874	2826.2394
414	905.3479	1083.4381	2835.3580
415	907.9660	1086.4900	2844.4830
416	910.5850	1089.5428	2853.6143
417	913.2052	1092.5967	2862.7521
418	915.8264	1095.6517	2871.8962
419	918.4486	1098.7077	2881.0467
420	921.0718	1101.7647	2890.2035
421	923.6961	1104.8228	2899.3666
422	926.3214	1107.8819	2908.5360
423	928.9478	1110.9420	2917.7117
424	931.5751	1114.0031	2926.8938
425	934.2035	1117.0653	2936.0820
426	936.8329	1120.1285	2945.2766
427	939.4633	1123.1927	2954.4774
428	942.0948	1126.2579	2963.6844
429	944.7272	1129.3242	2972.8976
430	947.3607	1132.3914	2982.1171
431	949.9952	1135.4597	2991.3428
432	952.6307	1138.5290	3000.5746
433	955.2672	1141.5993	3009.8126
434	957.9047	1144.6705	3019.0568
435	960.5431	1147.7428	3028.3071
436	963.1826	1150.8161	3037.5636
437	965.8231	1153.8904	3046.8261
438	968.4646	1156.9657	3056.0948
439	971.1071	1160.0419	3065.3696
440	973.7505	1163.1192	3074.6505
441	976.3949	1166.1974	3083.9374
442	979.0404	1169.2766	3093.2305
443	981.6868	1172.3568	3102.5295
444	984.3342	1175.4380	3111.8346
445	986.9825	1178.5202	3121.1458
446	989.6318	1181.6033	3130.4629
447	992.2822	1184.6875	3139.7861
448	994.9334	1187.7725	3149.1152
449	997.5857	1190.8586	3158.4504
450	1000.2389	1193.9456	3167.7915
451	1002.8931	1197.0336	3177.1385
452	1005.5482	1200.1226	3186.4915
453	1008.2043	1203.2125	3195.8505
454	1010.8614	1206.3033	3205.2154
455	1013.5194	1209.3952	3214.5862
456	1016.1783	1212.4880	3223.9629
457	1018.8382	1215.5817	3233.3455
458	1021.4991	1218.6764	3242.7339
459	1024.1609	1221.7720	3252.1283
460	1026.8237	1224.8686	3261.5284
461	1029.4874	1227.9661	3270.9345
462	1032.1520	1231.0646	3280.3464
463	1034.8176	1234.1640	3289.7641
464	1037.4841	1237.2643	3299.1876
465	1040.1516	1240.3656	3308.6169
466	1042.8200	1243.4678	3318.0521
467	1045.4893	1246.5710	3327.4930
468	1048.1595	1249.6750	3336.9396
469	1050.8307	1252.7801	3346.3921
470	1053.5028	1255.8860	3355.8503
471	1056.1758	1258.9928	3365.3142
472	1058.8498	1262.1006	3374.7839
473	1061.5246	1265.2093	3384.2592
474	1064.2004	1268.3189	3393.7403
475	1066.8771	1271.4295	3403.2271
476	1069.5547	1274.5409	3412.7196
477	1072.2332	1277.6533	3422.2177
478	1074.9126	1280.7665	3431.7216
479	1077.5930	1283.8807	3441.2310
480	1080.2742	1286.9958	3450.7462
481	1082.9564	1290.1118	3460.2669
482	1085.6394	1293.2287	3469.7933
483	1088.3234	1296.3465	3479.3253
484	1091.0082	1299.4651	3488.8630
485	1093.6940	1302.5847	3498.4062
486	1096.3806	1305.7052	3507.9550
487	1099.0681	1308.8266	3517.5094
488	1101.7565	1311.9489	3527.0693
489	1104.4458	1315.0720	3536.6349
490	1107.1360	1318.1961	3546.2059
491	1109.8271	1321.3210	3555.7825
492	1112.5191	1324.4468	3565.3646
493	1115.2119	1327.5735	3574.9523
494	1117.9057	1330.7011	3584.5454
495	1120.6003	1333.8296	3594.1441
496	1123.2957	1336.9589	3603.7482
497	1125.9921	1340.0891	3613.3578
498	1128.6893	1343.2202	3622.9730
499	1131.3874	1346.3522	3632.5935
500	1134.0864	1349.4850	3642.2195
501	1136.7862	1352.6187	3651.8510
502	1139.4869	1355.7533	3661.4879
503	1142.1885	1358.8887	3671.1302
504	1144.8909	1362.0250	3680.7779
505	1147.5942	1365.1621	3690.4310
506	1150.2984	1368.3002	3700.0896
507	1153.0034	1371.4390	3709.7535
508	1155.7093	1374.5788	3719.4228
509	1158.4160	1377.7193	3729.0974
510	1161.1235	1380.8608	3738.7775
511	1163.8320	1384.0031	3748.4629
512	1166.5412	1387.1462	3758.1536
513	1169.2514	1390.2902	3767.8496
514	1171.9623	1393.4350	3777.5510
515	1174.6741	1396.5807	3787.2577
516	1177.3868	1399.7272	3796.9697
517	1180.1003	1402.8746	3806.6870
518	1182.8146	1406.0228	3816.4095
519	1185.5298	1409.1718	3826.1374
520	1188.2458	1412.3217	3835.8705
521	1190.9626	1415.4724	3845.6089
522	1193.6803	1418.6240	3855.3526
523	1196.3988	1421.7764	3865.1015
524	1199.1181	1424.9296	3874.8556
525	1201.8383	1428.0836	3884.6150
526	1204.5592	1431.2385	3894.3795
527	1207.2811	1434.3942	3904.1493
528	1210.0037	1437.5507	3913.9243
529	1212.7271	1440.7080	3923.7045
530	1215.4514	1443.8662	3933.4899
531	1218.1765	1447.0252	3943.2804
532	1220.9024	1450.1850	3953.0761
533	1223.6292	1453.3456	3962.8770
534	1226.3567	1456.5070	3972.6830
535	1229.0851	1459.6693	3982.4942
536	1231.8142	1462.8323	3992.3105
537	1234.5442	1465.9962	4002.1319
538	1237.2750	1469.1609	4011.9584
539	1240.0066	1472.3263	4021.7901
540	1242.7390	1475.4926	4031.6268
541	1245.4722	1478.6597	4041.4687
542	1248.2061	1481.8276	4051.3156
543	1250.9409	1484.9963	4061.1676
544	1253.6765	1488.1658	4071.0247
545	1256.4129	1491.3361	4080.8868
546	1259.1501	1494.5072	4090.7541
547	1261.8881	1497.6791	4100.6263
548	1264.6269	1500.8517	4110.5035
549	1267.3665	1504.0252	4120.3859
550	1270.1068	1507.1995	4130.2732
551	1272.8480	1510.3745	4140.1655
552	1275.5899	1513.5504	4150.0629
553	1278.3327	1516.7270	4159.9652
554	1281.0762	1519.9044	4169.8726
555	1283.8205	1523.0826	4179.7849
556	1286.5655	1526.2616	4189.7021
557	1289.3114	1529.4413	4199.6245
558	1292.0580	1532.6219	4209.5516
559	1294.8054	1535.8032	4219.4838
560	1297.5536	1538.9853	4229.4210
561	1300.3026	1542.1682	4239.3630
562	1303.0523	1545.3518	4249.3099
563	1305.8028	1548.5362	4259.2618
564	1308.5541	1551.7214	4269.2187
565	1311.3062	1554.9074	4279.1804
566	1314.0590	1558.0941	4289.1470
567	1316.8126	1561.2816	4299.1185
568	1319.5669	1564.4698	4309.0949
569	1322.3220	1567.6589	4319.0762
570	1325.0779	1570.8487	4329.0623
571	1327.8345	1574.0392	4339.0533
572	1330.5919	1577.2305	4349.0492
573	1333.3501	1580.4226	4359.0499
574	1336.1090	1583.6154	4369.0554
575	1338.8687	1586.8090	4379.0658
576	1341.6291	1590.0033	4389.0810
577	1344.3903	1593.1984	4399.1010
578	1347.1522	1596.3943	4409.1258
579	1349.9149	1599.5909	4419.1555
580	1352.6783	1602.7882	4429.1898
581	1355.4425	1605.9863	4439.2291
582	1358.2074	1609.1852	4449.2731
583	1360.9731	1612.3848	4459.3218
584	1363.7395	1615.5851	4469.3754
585	1366.5066	1618.7862	4479.4337
586	1369.2745	1621.9880	4489.4967
587	1372.0432	1625.1906	4499.5645
588	1374.8125	1628.3939	4509.6370
589	1377.5827	1631.5979	4519.7143
590	1380.3535	1634.8027	4529.7963
591	1383.1251	1638.0082	4539.8830
592	1385.8974	1641.2144	4549.9744
593	1388.6705	1644.4214	4560.0706
594	1391.4443	1647.6291	4570.1714
595	1394.2188	1650.8376	4580.2769
596	1396.9940	1654.0468	4590.3871

TABLE 23. (Continued)

n	log n!	n log n	n log² n	n	log n!	n log n	n log² n	n	log n!	n log n	n log² n	n	log n!	n log n	n log² n
597	1399.7700	1657.2567	4600.5020	654	1559.1662	1841.3878	5184.5706	711	1720.7210	2027.6793	5782.6769	768	1884.2611	2215.9574	6393.8376
598	1402.5467	1660.4673	4610.6215	655	1561.9824	1844.6380	5194.9459	712	1723.5735	2030.9657	5793.2892	769	1887.1470	2219.2773	6404.6710
599	1405.3241	1663.6787	4620.7457	656	1564.7993	1847.8889	5205.3254	713	1726.4265	2034.2528	5803.9055	770	1890.0335	2222.5978	6415.5081
600	1408.1023	1666.8907	4630.8746	657	1567.6169	1851.1404	5215.7092	714	1729.2802	2037.5405	5814.5257	771	1892.9205	2225.9189	6426.3489
601	1410.8811	1670.1035	4641.0081	658	1570.4351	1854.3926	5226.0973	715	1732.1346	2040.8288	5825.1500	772	1895.8082	2229.2405	6437.1935
602	1413.6608	1673.3171	4651.1463	659	1573.2540	1857.6455	5236.4897	716	1734.9895	2044.1177	5835.7783	773	1898.6963	2232.5627	6448.0419
603	1416.4411	1676.5313	4661.2891	660	1576.0735	1860.8990	5246.8863	717	1737.8450	2047.4072	5846.4105	774	1901.5851	2235.8855	6458.8940
604	1419.2221	1679.7463	4671.4365	661	1578.8938	1864.1532	5257.2875	718	1740.7011	2050.6973	5857.0468	775	1904.4744	2239.2088	6469.7498
605	1422.0039	1682.9620	4681.5886	662	1581.7146	1867.4080	5267.6927	719	1743.5578	2053.9881	5867.6870	776	1907.3642	2242.5327	6480.6094
606	1424.7863	1686.1784	4691.7452	663	1584.5361	1870.6635	5278.1022	720	1746.4152	2057.2794	5878.3312	777	1910.2547	2245.8571	6491.4727
607	1427.5695	1689.3955	4701.9065	664	1587.3583	1873.9196	5288.5161	721	1749.2731	2060.5713	5888.9794	778	1913.1456	2249.1821	6502.3396
608	1430.3534	1692.6134	4712.0724	665	1590.1811	1877.1764	5298.9341	722	1752.1316	2063.8638	5899.6316	779	1916.0372	2252.5077	6513.2103
609	1433.1380	1695.8319	4722.2429	666	1593.0046	1880.4338	5309.3564	723	1754.9908	2067.1570	5910.2877	780	1918.9293	2255.8338	6524.0847
610	1435.9234	1699.0512	4732.4180	667	1595.8287	1883.6919	5319.7830	724	1757.8505	2070.4507	5920.9478	781	1921.8219	2259.1604	6534.9628
611	1438.7094	1702.2712	4742.5977	668	1598.6535	1886.9507	5330.2138	725	1760.7109	2073.7450	5931.6118	782	1924.7151	2262.4877	6545.8446
612	1441.4962	1705.4919	4752.7819	669	1601.4789	1890.2101	5340.6489	726	1763.5718	2077.0400	5942.2797	783	1927.6089	2265.8154	6556.7301
613	1444.2836	1708.7133	4762.9707	670	1604.3050	1893.4701	5351.0881	727	1766.4333	2080.3355	5952.9517	784	1930.5032	2269.1438	6567.6193
614	1447.0718	1711.9354	4773.1641	671	1607.1317	1896.7308	5361.5317	728	1769.2955	2083.6316	5963.6275	785	1933.3981	2272.4727	6578.5122
615	1449.8607	1715.1582	4783.3620	672	1609.9591	1899.9921	5371.9794	729	1772.1582	2086.9283	5974.3073	786	1936.2935	2275.8021	6589.4088
616	1452.6503	1718.3817	4793.5645	673	1612.7871	1903.2541	5382.4313	730	1775.0215	2090.2257	5984.9910	787	1939.1895	2279.1321	6600.3090
617	1455.4405	1721.6059	4803.7715	674	1615.6158	1906.5168	5392.8875	731	1777.8854	2093.5236	5995.6786	788	1942.0860	2282.4626	6611.2129
618	1458.2315	1724.8309	4813.9831	675	1618.4451	1909.7800	5403.3478	732	1780.7499	2096.8221	6006.3701	789	1944.9831	2285.7937	6622.1205
619	1461.0232	1728.0565	4824.1992	676	1621.2750	1913.0440	5413.8124	733	1783.6150	2100.1212	6017.0656	790	1947.8807	2289.1254	6633.0317
620	1463.8156	1731.2828	4834.4198	677	1624.1056	1916.3085	5424.2812	734	1786.4807	2103.4209	6027.7650	791	1950.7789	2292.4576	6643.9467
621	1466.6087	1734.5099	4844.6450	678	1626.9368	1919.5737	5434.7542	735	1789.3470	2106.7212	6038.4683	792	1953.6776	2295.7903	6654.8652
622	1469.4025	1737.7376	4854.8746	679	1629.7687	1922.8396	5445.2313	736	1792.2139	2110.0221	6049.1754	793	1956.5769	2299.1236	6665.7875
623	1472.1970	1740.9660	4865.1088	680	1632.6012	1926.1060	5455.7125	737	1795.0814	2113.3235	6059.8865	794	1959.4767	2302.4575	6676.7134
624	1474.9922	1744.1952	4875.3475	681	1635.4344	1929.3732	5466.1981	738	1797.9494	2116.6256	6070.6015	795	1962.3771	2305.7918	6687.6429
625	1477.7880	1747.4250	4885.5906	682	1638.2681	1932.6409	5476.6877	739	1800.8181	2119.9282	6081.3203	796	1965.2780	2309.1268	6698.5761
626	1480.5846	1750.6555	4895.8383	683	1641.1026	1935.9093	5487.1815	740	1803.6873	2123.2314	6092.0430	797	1968.1794	2312.4623	6709.5129
627	1483.3819	1753.8867	4906.0905	684	1643.9376	1939.1784	5497.6794	741	1806.5571	2126.5353	6102.7697	798	1971.0814	2315.7983	6720.4534
628	1486.1798	1757.1187	4916.3470	685	1646.7733	1942.4480	5508.1816	742	1809.4275	2129.8397	6113.5002	799	1973.9840	2319.1349	6731.3975
629	1488.9785	1760.3512	4926.6082	686	1649.6096	1945.7183	5518.6878	743	1812.2985	2133.1447	6124.2345	800	1976.8871	2322.4720	6742.3452
630	1491.7778	1763.5845	4936.8737	687	1652.4466	1948.9893	5529.1982	744	1815.1701	2136.4503	6134.9727	801	1979.7907	2325.8096	6753.2965
631	1494.5779	1766.8185	4947.1437	688	1655.2842	1952.2608	5539.7128	745	1818.0422	2139.7564	6145.7148	802	1982.6949	2329.1478	6764.2515
632	1497.3786	1770.0532	4957.4182	689	1658.1224	1955.5330	5550.2314	746	1820.9150	2143.0631	6156.4608	803	1985.5996	2332.4866	6775.2100
633	1500.1800	1773.2885	4967.6971	690	1660.9612	1958.8059	5560.7542	747	1823.7883	2146.3705	6167.2105	804	1988.5049	2335.8258	6786.1722
634	1502.9821	1776.5246	4977.9804	691	1663.8007	1962.0793	5571.2811	748	1826.6622	2149.6784	6177.9642	805	1991.4106	2339.1657	6797.1380
635	1505.7849	1779.7613	4988.2682	692	1666.6408	1965.3534	5581.8122	749	1829.5367	2152.9869	6188.7216	806	1994.3170	2342.5060	6808.1074
636	1508.5883	1782.9987	4998.5604	693	1669.4816	1968.6281	5592.3474	750	1832.4117	2156.2959	6199.4829	807	1997.2239	2345.8469	6819.0804
637	1511.3924	1786.2368	5008.8571	694	1672.3229	1971.9035	5602.8866	751	1835.2874	2159.6056	6210.2480	808	2000.1313	2349.1884	6830.0569
638	1514.1973	1789.4756	5019.1581	695	1675.1649	1975.1794	5613.4299	752	1838.1636	2162.9158	6221.0170	809	2003.0392	2352.5303	6841.0371
639	1517.0028	1792.7150	5029.4636	696	1678.0075	1978.4560	5623.9774	753	1841.0404	2166.2266	6231.7898	810	2005.9477	2355.8729	6852.0209
640	1519.8089	1795.9552	5039.7734	697	1680.8507	1981.7332	5634.5289	754	1843.9178	2169.5380	6242.5664	811	2008.8567	2359.2159	6863.0082
641	1522.6158	1799.1960	5050.0877	698	1683.6946	1985.0111	5645.0845	755	1846.7957	2172.8499	6253.3469	812	2011.7663	2362.5595	6873.9992
642	1525.4233	1802.4375	5060.4064	699	1686.5391	1988.2895	5655.6442	756	1849.6742	2176.1625	6264.1311	813	2014.6764	2365.9036	6884.9937
643	1528.2316	1805.6796	5070.7294	700	1689.3842	1991.5686	5666.2079	757	1852.5533	2179.4756	6274.9191	814	2017.5870	2369.2483	6895.9918
644	1531.0404	1808.9225	5081.0568	701	1692.2299	1994.8483	5676.7758	758	1855.4330	2182.7892	6285.7110	815	2020.4982	2372.5934	6906.9935
645	1533.8500	1812.1660	5091.3886	702	1695.0762	1998.1286	5687.3477	759	1858.3132	2186.1035	6296.5066	816	2023.4099	2375.9391	6917.9987
646	1536.6602	1815.4102	5101.7248	703	1697.9232	2001.4096	5697.9236	760	1861.1941	2189.4183	6307.3060	817	2026.3221	2379.2854	6929.0074
647	1539.4711	1818.6551	5112.0653	704	1700.7708	2004.6911	5708.5037	761	1864.0754	2192.7337	6318.1093	818	2029.2348	2382.6322	6940.0198
648	1542.2827	1821.9006	5122.4102	705	1703.6189	2007.9733	5719.0878	762	1866.9574	2196.0497	6328.9163	819	2032.1481	2385.9795	6951.0357
649	1545.0950	1825.1468	5132.7594	706	1706.4678	2011.2561	5729.6758	763	1869.8399	2199.3662	6339.7271	820	2035.0619	2389.3273	6962.0551
650	1547.9079	1828.3937	5143.1130	707	1709.3172	2014.5395	5740.2680	764	1872.7230	2202.6833	6350.5416	821	2037.9763	2392.6757	6973.0781
651	1550.7215	1831.6412	5153.4709	708	1712.1672	2017.8235	5750.8642	765	1875.6067	2206.0010	6361.3600	822	2040.8911	2396.0246	6984.1047
652	1553.5357	1834.8894	5163.8331	709	1715.0179	2021.1082	5761.4644	766	1878.4909	2209.3192	6372.1821	823	2043.8065	2399.3741	6995.1347
653	1556.3506	1838.1383	5174.1993	710	1717.8691	2024.3934	5772.0686	767	1881.3757	2212.6380	6383.0079	824	2046.7225	2402.7240	7006.1683

TABLE 23. (Continued)

n	log n!	n log n	$n \log^2 n$	n	log n!	n log n	$n \log^2 n$
825	2049.6389	2406.0745	7017.2094	882	2216.7274	2597.9033	7652.0425
826	2052.5559	2409.4255	7028.2462	883	2219.6734	2601.2833	7663.2784
827	2055.4734	2412.7770	7039.2903	884	2222.6198	2604.6638	7674.5175
828	2058.3914	2416.1291	7050.3380	885	2225.5668	2608.0448	7685.7600
829	2061.3100	2419.4817	7061.3893	886	2228.5142	2611.4263	7697.0059
830	2064.2291	2422.8348	7072.4440	887	2231.4621	2614.8082	7708.2549
831	2067.1487	2426.1884	7083.5023	888	2234.4105	2618.1907	7719.5074
832	2070.0688	2429.5426	7094.5640	889	2237.3594	2621.5736	7730.7632
833	2072.9894	2432.8973	7105.6293	890	2240.3088	2624.9571	7742.0222
834	2075.9106	2436.2525	7116.6980	891	2243.2587	2628.3410	7753.2846
835	2078.8323	2439.6082	7127.7703	892	2246.2091	2631.7254	7764.5502
836	2081.7545	2442.9644	7138.8460	893	2249.1599	2635.1104	7775.8192
837	2084.6772	2446.3212	7149.9252	894	2252.1113	2638.4957	7787.0914
838	2087.6005	2449.6785	7161.0079	895	2255.0631	2641.8816	7798.3670
839	2090.5242	2453.0363	7172.0942	896	2258.0154	2645.2680	7809.6458
840	2093.4485	2456.3946	7183.1838	897	2260.9682	2648.6548	7820.9279
841	2096.3733	2459.7534	7194.2770	898	2263.9214	2652.0421	7832.2133
842	2099.2986	2463.1128	7205.3736	899	2266.8752	2655.4300	7843.5020
843	2102.2244	2466.4726	7216.4736	900	2269.8295	2658.8182	7854.7939
844	2105.1508	2469.8330	7227.5772	901	2272.7842	2662.2070	7866.0891
845	2108.0776	2473.1939	7238.6842	902	2275.7394	2665.5963	7877.3876
846	2111.0050	2476.5553	7249.7946	903	2278.6951	2668.9860	7888.6893
847	2113.9329	2479.9172	7260.9086	904	2281.6512	2672.3763	7899.9943
848	2116.8613	2483.2797	7272.0259	905	2284.6079	2675.7669	7911.3026
849	2119.7902	2486.6426	7283.1467	906	2287.5650	2679.1581	7922.6141
850	2122.7196	2490.0061	7294.2709	907	2290.5226	2682.5498	7933.9288
851	2125.6495	2493.3700	7305.3986	908	2293.4807	2685.9419	7945.2468
852	2128.5800	2496.7345	7316.5297	909	2296.4393	2689.3346	7956.5681
853	2131.5109	2500.0995	7327.6642	910	2299.3983	2692.7277	7967.8926
854	2134.4424	2503.4650	7338.8022	911	2302.3578	2696.1212	7979.2203
855	2137.3744	2506.8310	7349.9436	912	2305.3178	2699.5153	7990.5513
856	2140.3068	2510.1975	7361.0884	913	2308.2783	2702.9098	8001.8855
857	2143.2398	2513.5645	7372.2366	914	2311.2393	2706.3048	8013.2230
858	2146.1733	2516.9321	7383.3882	915	2314.2007	2709.7003	8024.5636
859	2149.1073	2520.3001	7394.5433	916	2317.1626	2713.0963	8035.9075
860	2152.0418	2523.6686	7405.7018	917	2320.1249	2716.4927	8047.2546
861	2154.9768	2527.0377	7416.8636	918	2323.0878	2719.8896	8058.6049
862	2157.9123	2530.4073	7428.0289	919	2326.0511	2723.2869	8069.9584
863	2160.8483	2533.7773	7439.1975	920	2329.0149	2726.6848	8081.3152
864	2163.7848	2537.1479	7450.3696	921	2331.9792	2730.0831	8092.6752
865	2166.7218	2540.5189	7461.5450	922	2334.9439	2733.4819	8104.0383
866	2169.6594	2543.8905	7472.7238	923	2337.9091	2736.8812	8115.4047
867	2172.5974	2547.2625	7483.9060	924	2340.8748	2740.2809	8126.7742
868	2175.5359	2550.6351	7495.0916	925	2343.8409	2743.6811	8138.1470
869	2178.4749	2554.0082	7506.2805	926	2346.8075	2747.0818	8149.5229
870	2181.4144	2557.3817	7517.4728	927	2349.7746	2750.4829	8160.9021
871	2184.3544	2560.7558	7528.6686	928	2352.7421	2753.8845	8172.2844
872	2187.2950	2564.1304	7539.8676	929	2355.7101	2757.2866	8183.6699
873	2190.2360	2567.5054	7551.0700	930	2358.6786	2760.6891	8195.0586
874	2193.1775	2570.8810	7562.2758	931	2361.6476	2764.0921	8206.4504
875	2196.1195	2574.2570	7573.4849	932	2364.6170	2767.4956	8217.8455
876	2199.0620	2577.6336	7584.6974	933	2367.5869	2770.8996	8229.2438
877	2202.0050	2581.0106	7595.9133	934	2370.5572	2774.3040	8240.6451
878	2204.9485	2584.3882	7607.1324	935	2373.5280	2777.7088	8252.0497
879	2207.8925	2587.7662	7618.3549	936	2376.4993	2781.1142	8263.4574
880	2210.8370	2591.1447	7629.5808	937	2379.4710	2784.5200	8274.8683
881	2213.7820	2594.5238	7640.8100	938	2382.4433	2787.9262	8286.2823

n	log n!	n log n	$n \log^2 n$	n	log n!	n log n	$n \log^2 n$
939	2385.4159	2791.3330	8297.6995	995	2552.6090	2982.8340	8942.0084
940	2388.3890	2794.7402	8309.1199	996	2555.6072	2986.2063	8953.6007
941	2391.3626	2798.1478	8320.5433	997	2558.6059	2989.6991	8965.1960
942	2394.3367	2801.5559	8331.9700	998	2561.6051	2993.1323	8976.7944
943	2397.3112	2804.9645	8343.3998	999	2564.6046	2996.5629	8988.3956
944	2400.2862	2808.3735	8354.8326	1000	2567.6046	3000.0000	9000.0000
945	2403.2616	2811.7831	8366.2687	1001	2570.6051	3003.4345	9011.6072
946	2406.2375	2815.1930	8377.7079	1002	2573.6059	3006.8694	9023.2174
947	2409.2138	2818.6034	8389.1503	1003	2576.6072	3010.3048	9034.8307
948	2412.1906	2822.0143	8400.5957	1004	2579.6090	3013.7406	9046.4469
949	2415.1679	2825.4256	8412.0442	1005	2582.6111	3017.1769	9058.0660
950	2418.1456	2828.8374	8423.4960	1006	2585.6137	3020.6136	9069.6881
951	2421.1238	2832.2497	8434.9507	1007	2588.6168	3024.0507	9081.3132
952	2424.1024	2835.6624	8446.4087	1008	2591.6202	3027.4882	9092.9413
953	2427.0815	2839.0755	8457.8698	1009	2594.6241	3030.9262	9104.5724
954	2430.0611	2842.4891	8469.3339	1010	2597.6284	3034.3646	9116.2063
955	2433.0411	2845.9032	8480.8011	1011	2600.6332	3037.8034	9127.8433
956	2436.0216	2849.3177	8492.2715	1012	2603.6384	3041.2427	9139.4832
957	2439.0025	2852.7327	8503.7450	1013	2606.6440	3044.6823	9151.1260
958	2441.9838	2856.1481	8515.2216	1014	2609.6500	3048.1225	9162.7719
959	2444.9657	2859.5640	8526.7012	1015	2612.6565	3051.5630	9174.4205
960	2447.9479	2862.9804	8538.1840	1016	2615.6634	3055.0040	9186.0723
961	2450.9307	2866.3972	8549.6698	1017	2618.6707	3058.4454	9197.7269
962	2453.9138	2869.8144	8561.1588	1018	2621.6784	3061.8872	9209.3845
963	2456.8975	2873.2321	8572.6508	1019	2624.6866	3065.3295	9221.0450
964	2459.8815	2876.6502	8584.1459	1020	2627.6952	3068.7722	9232.7084
965	2462.8661	2880.0688	8595.6442	1021	2630.7043	3072.2153	9244.3749
966	2465.8510	2883.4879	8607.1454	1022	2633.7137	3075.6588	9256.0441
967	2468.8365	2886.9074	8618.6498	1023	2636.7236	3079.1028	9267.7163
968	2471.8223	2890.3273	8630.1571	1024	2639.7339	3082.5471	9279.3915
969	2474.8087	2893.7477	8641.6676	1025	2642.7446	3085.9919	9291.0696
970	2477.7954	2897.1686	8653.1812	1026	2645.7557	3089.4372	9302.7506
971	2480.7827	2900.5898	8664.6978	1027	2648.7673	3092.8828	9314.4346
972	2483.7703	2904.0116	8676.2174	1028	2651.7793	3096.3289	9326.1213
973	2486.7584	2907.4338	8687.7402	1029	2654.7917	3099.7754	9337.8110
974	2489.7470	2910.8564	8699.2660	1030	2657.8046	3103.2223	9349.5038
975	2492.7360	2914.2795	8710.7948	1031	2660.8178	3106.6697	9361.1993
976	2495.7254	2917.7030	8722.3267	1032	2663.8315	3110.1174	9372.8977
977	2498.7153	2921.1270	8733.8617	1033	2666.8456	3113.5656	9384.5991
978	2501.7057	2924.5514	8745.3997	1034	2669.8601	3117.0142	9396.3033
979	2504.6965	2927.9762	8756.9407	1035	2672.8751	3120.4633	9408.0105
980	2507.6877	2931.4016	8768.4847	1036	2675.8904	3123.9127	9419.7206
981	2510.6794	2934.8273	8780.0319	1037	2678.9062	3127.3625	9431.4336
982	2513.6715	2938.2535	8791.5819	1038	2681.9224	3130.8128	9443.1494
983	2516.6640	2941.6801	8803.1351	1039	2684.9390	3134.2635	9454.8682
984	2519.6570	2945.1071	8814.6913	1040	2687.9560	3137.7147	9466.5897
985	2522.6505	2948.5347	8826.2503	1041	2690.9735	3141.1662	9478.3142
986	2525.6443	2951.9626	8837.8127	1042	2693.9914	3144.6181	9490.0416
987	2528.6386	2955.3910	8849.3781	1043	2697.0096	3148.0705	9501.7719
988	2531.6334	2958.8199	8860.9463	1044	2700.0284	3151.5233	9513.5050
989	2534.6286	2962.2491	8872.5176	1045	2703.0475	3154.9765	9525.2410
990	2537.6242	2965.6788	8884.0918	1046	2706.0670	3158.4301	9536.9799
991	2540.6203	2969.1090	8895.6692	1047	2709.0869	3161.8842	9548.7216
992	2543.6168	2972.5396	8907.2495	1048	2712.1073	3165.3386	9560.4662
993	2546.6138	2975.9706	8918.8328	1049	2715.1281	3168.7935	9572.2136
994	2549.6111	2979.4020	8930.4191	1050	2718.1493	3172.2487	9583.9640

TABLE 24. THE DIVERSITY OF SPECIES, $\overline{d}$, CHARACTERISTIC OF MacARTHUR'S MODEL FOR VARIOUS NUMBERS OF HYPOTHETICAL SPECIES, s'*

s'	$\overline{d}$	s'	$\overline{d}$	s'	$\overline{d}$	s'	$\overline{d}$
1	0.0000	51	5.0941	102	6.0792	205	7.078
2	0.8113	52	5.1215	104	6.1069	210	7.112
3	1.2997	53	5.1485	106	6.1341	215	7.146
4	1.6556	54	5.1749	108	6.1608	220	7.179
5	1.9374	55	5.2009	110	6.1870	225	7.211
6	2.1712	56	5.2264	112	6.2128	230	7.243
7	2.3714	57	5.2515	114	6.2380	235	7.274
8	2.5465	58	5.2761	116	6.2629	240	7.304
9	2.7022	59	5.3004	118	6.2873	245	7.334
10	2.8425	60	5.3242	120	6.3113	250	7.363
11	2.9701	61	5.3476	122	6.3350	255	7.391
12	3.0872	62	5.3707	124	6.3582	260	7.419
13	3.1954	63	5.3934	126	6.3811	265	7.446
14	3.2960	64	5.4157	128	6.4036	270	7.473
15	3.3899	65	5.4378	130	6.4258	275	7.500
16	3.4780	66	5.4594	132	6.4476	280	7.52
17	3.5611	67	5.4808	134	6.4691	285	7.551
18	3.6395	68	5.5018	136	6.4903	290	7.576
19	3.7139	69	5.5226	138	6.5112	295	7.600
20	3.7846	70	5.5430	140	6.5318	300	7.62
21	3.8520	71	5.5632	142	6.5521	310	7.67
22	3.9163	72	5.5830	144	6.5721	320	7.71
23	3.9779	73	5.6027	146	6.5919	330	7.76
24	4.0369	74	5.6220	148	6.6114	340	7.80
25	4.0937	75	5.6411	150	6.6306	350	7.84
26	4.1482	76	5.6599	152	6.6495	360	7.88
27	4.2008	77	5.6785	154	6.6683	370	7.92
28	4.2515	78	5.6969	156	6.6867	380	7.96
29	4.3004	79	5.7150	158	6.7050	390	8.00
30	4.3478	80	5.7329	160	6.7230	400	8.03
31	4.3936	81	5.7506	162	6.7408	410	8.07
32	4.4381	82	5.7681	164	6.7584	420	8.10
33	4.4812	83	5.7853	166	6.7757	430	8.14
34	4.5230	84	5.8024	168	6.7929	440	8.17
35	4.5637	85	5.8192	170	6.8099	450	8.20
36	4.6032	86	5.8359	172	6.8266	460	8.23
37	4.6417	87	5.8524	174	6.8432	470	8.27
38	4.6792	88	5.8687	176	6.8596	480	8.30
39	4.7157	89	5.8848	178	6.8758	490	8.33
40	4.7513	90	5.9007	180	6.8918	500	8.35
41	4.7861	91	5.9164	182	6.9076	550	8.49
42	4.8200	92	5.9320	184	6.9233	600	8.62
43	4.8532	93	5.9474	186	6.9388	650	8.73
44	4.8856	94	5.9627	188	6.9541	700	8.84
45	4.9173	95	5.9778	190	6.9693	750	8.94
46	4.9483	96	5.9927	192	6.9843	800	9.03
47	4.9787	97	6.0075	194	6.9992	850	9.12
48	5.0084	98	6.0221	196	7.0139	900	9.20
49	5.0375	99	6.0366	198	7.0284	950	9.28
50	5.0661	100	6.0510	200	7.0429	1000	9.35

*The data in this table are reproduced, with permission, from Lloyd and Ghelardi

7.7 Literature Cited

Allan, J.D. 1984. Hypothesis testing in ecological studies of aquatic insects. In: V.H. Resh and D.M. Rosenberg, eds. The ecology of aquatic insects. Praeger Scientific, NY. pp. 484-507.

Beck, W.M., Jr. 1954. Studies in stream pollution biology: I. A simplified ecological classification of organisms. J. Fla. Acad. Sci. 17(4):211-227.

Beck, W.M., Jr. 1955. Suggested method for reporting biotic data. Sewage Ind. Wastes 27(10):1193-1197.

Boesch, D.F. 1977. Application of numerical classification in ecological investigations of water pollution. EPA-600/3-77-033. U.S. Environmental Protection Agency, Corvallis, OR.

Bray, J.R. and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27:325.

Cairns, J., Jr. and K.L. Dickson. 1971. A simple method for the biological assessment of the effects of waste discharges on aquatic bottom - dwelling organisms. J. Wat. Pollut. Control Fed. 43(5):755-772.

Chutter, F.M. 1972. An empirical biotic index of the quality of water in South African streams and rivers. Wat. Res. 6:19-30.

Clements, W.H., D.S. Cherry, and J. Cairns, Jr. 1988. Impact of heavy metals on insect communities in streams; a comparison of observational and experimental results. Can. J. Fish. Aquat. Sci. 45:2017-2025.

Conover, W.J. 1980. Practical nonparametric statistics. Second Edition. John Wiley and Sons, NY.

Cochran, W.G. 1952. The χ^2 test of goodness of fit. Ann. Math. Statistics. 23:315-345.

Courtemanch, D.L. and S.P. Davies. 1987. A coefficient of community loss to assess detrimental change in aquatic communities. Wat. Res. 21(2):217-222.

Cummins, K.W. 1973. Trophic relations of aquatic insects. Ann. Rev. Entomol. 18:183-206.

Dixon, W.J. and E.J Massey, Jr. 1983. Introduction to statistical analysis. (Fourth Edition), McGraw Hill, NY.

Ferrington, L.C. 1987. Collection and identification of floating exuviae of Chironomidae for use in studies of surface water quality. SOP No. FW 130A. U.S. Environmental Protection Agency, Region VII, Kansas City, KS.

Freund, J.E. 1986. Statistics a first course. (Fourth Edition) Prentice-

Hall, Inc., Eanglewood Cliffs, NJ. 557 pp.

Gauch, H.G. and R.H. Whittaker. 1972. Comparison of ordination techniques. Ecology 53(5):868-875.

Gaufin, A.R. and C.M. Tarzwell. 1956. Aquatic macroinvertebrate communities as indicators of organic pollution in Lytle Creek. Sew. Ind. Wastes 28:906-924.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, NY. 257 pp.

Hargrove, B.T. 1972. Aerobic decomposition of sediment and detritus as a function of particle surface area and organic content. Limnol. Oceanogr. 17:583-596.

Harkins, R.D. and R.E. Austin. 1973. Reduction and evaluation of biological data. J. Wat. Pollut. Control Fed. 45(7):1606-1611.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Tech. Bull. No. 100, Department of Natural Resources, Madison, Wisconsin. 15 pp.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomol. 20:31-39.

Hilsenhoff, W.L. 1988a. Rapid field assessment of organic pollution with a family-level biotic index. J. N. Am. Benthol. Soc. 7(1):65-68.

Hilsenhoff, W.L. 1988b. Seasonal correction factors for the biotic index. Great Lakes Entomol. 21(1):9-13.

Howmiller, R.P. and M.A. Scott. 1977. An environmental index based on relatiave abundance of oligochaete species. J. Wat. Pollut. Control Fed. 49(5):809-815.

Ingram, W.M. and A.F. Bartsch. 1960. Graphic expression of biological data in water pollution reports. J. Wat. Pollut. Control Fed. 32(3):297-310.

Jaccard, P. 1912. The distribution of flora in an alpine zone. New Phytol. 11:37.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6(6):12-27.

Kolkwitz, R. and M. Marsson. 1908. Oekologie der pflanzlichen Saprobien. Ber. Dtsch. Bot. Ges. 26(9):505

Lenat, D.R. 1983. Chironomid taxa richness: Natural variation and use in pollution assessment. Freshwat. Invertebr. Biol. 2(4):192-198.

Lewis, P.A. 1974. Taxonomy and ecology of <u>Stenonema</u> mayflies (Heptageniidae:

Ephemeroptera). EPA-670/4-74-006. Environmental Monitoring Series. U.S. Environmental Protection Agency, Methods Development and Quality Assurance Research Laboratory, National Environmental Research Center, Office of Research and Development, Cincinnati, Ohio 45268.

Lloyd, M. and R.J. Ghelardi. 1964. A table for calculating the "equitability" component of species diversity. J. Anim. Ecol. 33:217-225.

Lloyd, M., J.H. Zar, and J.R. Karr. 1968. On the calculation of information - theoretical measures of diversity. Am. Midl. Nat. 79 (2):257-272.

Margalef, D.R. 1957. Information theory in ecology. Yearbook of the Society for General Systems Research. 3:36-71.

MacArthur, R.H.J. 1957. On the relative abundance of bird species. Proc. Nat. Acad. Sci. Wash. 43:293-295.

McIntosh, R.P. 2967. An index of diversity and the relation of certain concepts to diversity. Ecol. 48(3):392-404.

Merritt, R.W. and K.W. Cummins (eds.). 1984. An introduction to the aquatic insects of North America (Second edition). Kendall/Hunt Publ. Co., Dubuque, IA.

Milbrink, G. 1983. An improved environmental index based on the relative abundance of oligochaete species. Hydrobiologia 102:89-97.

Moriseta, M. 1959. Measuring of interspecific association and similarity between communities. Memoirs Faculty Sci, Kyoshu Univ. Ser. E. Biol. 3-65.

Mount, D.J., N.A. Thomas, T.J. Norberg, M.T. Barbour, T.H. Roush, and W.F. Brandes. 1984. Effluent and ambient toxicity testing and instream community response on the Ottawa River, Lima, Ohio. EPA-600/3-84-080. U.S. Environmental Protection Agency, Environmental Research Laboratory, Duluth, MN.

Odum, E. P. (ed.). 1971. Fundamentals of Ecology. W. B. Saunders Company, Philadelphia, London, Toronto.

Ohio EPA, 1987. Biological criteria for the protection of aquatic life. Vol. II. Users manual for biological field assessment of Ohio surface water. Division of Water Quality Monitoring and Assessment, Surface Water Monitoring Section, Ohio EPA, Columbus, Ohio.

Ohio EPA, 1989. Biological criteria for the protection of aquatic life: Vol. III. Standardized biological field sampling and laboratoary methods for assessing fish and macroinvertebrate communities. Division of Water Quality Planning and Assessment Division, Surface Water Section, Ohio Environmental Protection Agency, Columbus, Ohio.

Patrick, R. 1950. Biological measure of stream conditions. Sewage Ind. Wastes 22(7):926-938.

Perkins, J.L. 1983. Bioassay evaluation of diversity and community comparison indexes. J. Wat. Pollut. Control Fed. 55:522-530.

Pinkham, C.F.A. and J.B. Pearson. 1976. Applications of a new coefficient of similarity to pollution surveys. J. Wat. Pollut. Control Fed. 48:717-723.

Plafkin, J.L., M.T.Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. EPA/440/4-89-001. U.S. Environmental Protection Agency, Assessment and Watershed Protection Division, Washington, DC 20460.

Rbeni, C.F., S.P. Davies, and K.E. Gibbs. 1985. Benthic invertebrate response to pollution abatement: Structural changes and functional implications. Wat. Resources Bull. 21(3):489-497.

Reynoldson, T.B., D.W. Schloesser, and B.A. Manny. 1989. Development of a benthic invertebrate objective for mesotrophic great lakes waters. J. Great Lakes Res. 15(4):669-686.

Richardson, R.E. 1928. The bottom fauna of the middle Illinois River 1913, 1925; its distribution, abundance, valuation, and index value in the study of stream pollution. Bull. Ill. State Nat. Hist. Sur. 17(12). 86 pp.

Schwenneker, B.W. and R.A. Hellenthal. 1984. Sampling considerations in using stream insects for monitoring water quality. Environ. Entomol. 13:741-750.

Shackleford, B. 1988. Rapid bioassessments of lotic macroinvertebrate communities: Biocriteria development. Arkansas Dept. Pollut. Control and Ecol., Little Rock, AR.

Shannon, C.E. and W. Weaver. 1963. The mathematical theory of communication. Univ. Illinois Press, Urbana, IL.

Snedecor, G.W. and W.G. Cochran, 1980. Statistical methods. Iowa State University Press, Ames, IA.

USEPA. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA-670/4-73-001. U.S. Environmental Protection Agency, Analytical Quality Control Laboratory, National Environmental Research Center, Office of Research and Development, Cincinnati, OH 45268.

USEPA. 1983. Technical Support manual: Water-body surveys and assessments for conducting use attainability analyses. U.S. Environmental Protection Agency, Washington, DC. 20460.

USEPA. 1989. Short-term methods for estimating the chronic toxicity of effluents and receiving water to freshwater organisms. EPA/600/4-89/001. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Cincinnati, Ohio 45268.

Washington, H.G. 1984. Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. Water Res. 18(6):653-694.

Weber, C.I. 1973. Biological monitoring of its aquatic environment by the Environmental Protection Agency. In: Biological methods for the assessment of water quality. ASTM STP 528. American Society for Testing and Materials, Philadelphia, PA. pp. 46-60.

Whittaker, R.H. 1952. A study of summer foliage insect communities in the Great Smoky Mountains. Ecol. Monogr. 22:6.

Wilhm, J.L. and T.C. Dorris. 1968. Biological parameters for water quality criteria. Bioscience 18:477-481.

Wilhm, J.L. 1970. Range of diversity index in benthic macroinvertebrate populations. J. Wat. Pollut. Control Fed. 42(5):R221-R224.

Winget, R.N. and F.A. Mangum. 1979. Biotic condition index. Integrated biological, physical, and chemical stream parameters for management. Intermountain Region, U.S. Dept. Agriculture, Forest Service, Ogden, Utah.

Winner, R.W., M.W. Boesel, and M.P. Farrell. 1980. Insect community structure as an index of heavy-metal pollution in lotic ecosystems. Can. J. Fish. Aquat. Sci. 37:647-655.

Wurtz, C.B. 1955. Stream biota and stream pollution. Sewage Ind. Wastes 27(11):1270-1278.

Zar, J.H. 1984. Biostatistical analysis. (Second Edition). Prentice-Hall, Inc., Englewood cliffs, NJ. 718 pp.